“Determining allometric relationships within tree species for a quantitative understanding of forest-atmosphere water fluxes coupled with remote-sensing-based methods for determining forest structure at an individual-tree scale”

HONORS RESEARCH DISTINCTION THESIS

Presented in Partial Fulfillment of the Requirements for Graduating with Honors Research Distinction in the College of Engineering of The Ohio State University

By

Kevin A. Meyer

Undergraduate Program in Environmental Engineering

The Ohio State University

2011

Honors Research Distinction Committee:

Dr. Gil Bohrer, Advisor
Dr. John Lenhart

Chapter 1 Abstract

Transpiration from forests contributes to both regional and global hydrological cycles. Rates of transpiration depend on a tree's physiology and its surrounding environmental factors. Many of these factors control water stress levels that in turn affect transpiration rates. By studying tree architecture and allometric relationships within that architecture, we can begin to quantify the 'pipe network' of a tree (i.e. its branching network). These relationships can produce parameters that define the hydrodynamic framework of a specific tree species. The framework can help further define the water stresses a tree experiences within its 'pipe network'. This information can be incorporated into hydrological models to enhance their predictive capabilities of forest-level transpiration. The models can consider changes in forest composition because each tree will have a different representative hydrodynamic framework. As forests change, the corresponding change in transpiration will be better incorporated into the models. Our results support that such hydrodynamic models can be developed. Virtual trees can be created that accurately depict a specific tree species. This is possible because many allometric relationships exist within a single tree, between trees of the same species, and between trees of different species.

Chapter 2 Abstract

Atmosphere-biosphere interactions involve exchanges of water, heat and carbon dioxide between the atmosphere and forest ecosystems. This exchange is important in contributing to weather and climate change. Current land-surface models do not resolve the effects of canopy structure change or incorporate 'real' forest heterogeneity as input parameters. The primary goal of this work is to provide a remote-sensing-based method to determine forest structure at an individual-tree scale. This will be done by: (1) parameterization of empirical allometric relationships that

govern the scaling of tree-crown hydraulic structure with tree size, species, and relative location in the canopy, and (2) tree-type classification of high-resolution airborne images combining visible and LIDAR. To determine allometric relationships, measurements of structural-hydraulic characteristics of trees, such as stem diameter and taper, branch diameter, splitting patterns, and leaf distribution were taken at existing experimental plots. 2010 census data containing tree species, location, and diameter at breast height (DBH) was obtained to provide additional forest composition information. Remote sensing data was provided by the National Agricultural Imagery Program. Aerial LIDAR data supplemented these images. Species-specific allometric equations were generated relating tree height and crown diameter to DBH, which will be input parameters to existing hydrological models. Using ENVI 4.5, an image analysis program, we will visually attempt to classify the heterogeneity of a forest. This will be coupled with the LIDAR data to provide additional classification schemes. The validity of our ENVI results will be compared to the 2010 census. Combining remote-sensing-determined composition with allometric relationships will allow simulations using new generation hydraulic models, which simulate water flow through the forest at the individual-tree scale and strengthen the predictive capabilities of existing hydrological models, and thus producing better predictions of atmosphere-biosphere interactions.

Acknowledgements

This work was conducted in large part thanks to a Biosphere-Atmosphere Research and Training (BART) summer REU fellowship from the University of Michigan's Biological Station (UMBS) to Kevin Meyer funded by NSF REU grant#AGS-0851421 to M.A. Carroll and D. Karowe. LIDAR data was provided through an NSF - National Center for Airborne Laser Mapping (NCALM) graduate seed award to Brady Hardiman, EEOB, OSU. Tree census data and site knowledge and support was provided by Chris Vogel and the staff of the UMBS and funded through the U.S. Department of Energy's Office of Science (BER) through the Midwestern Regional Center of the National Institute for Global Environmental Change under Cooperative Agreements No. DE-FC03-90ER610100, and the Midwestern Regional Center of the National Institute for Climatic Change Research at Michigan Technological University, under Award No. DE-FC02-06ER64158. Additional funding for the work was provided by NSF grant # DEB-0911461 and USDA-National Institute for Food & Agriculture (NIFA)- Air Quality Grant #2010-65112-20564 to Gil Bohrer. Furthermore, I would like express my sincere gratitutdue to Kyle Maurer (OSU CEEGS) for assisting with MATLAB programming and to Jenn Coury for assisting with a multitude of data collection.

Vita

June 2006……Lakota East High School

June 2011……B.S. Environmental Engineering, The Ohio State University

Fields of Study

Major Field: Environmental Engineering

Table of Contents

List of Tables

List of Figures

Chapter 1: Parameterization of empirical allometric relationships that govern the scaling of tree-crown hydraulic structure with tree size, species, and relative location in the canopy

1.0. Introduction

The hydrologic cycle is one of many integral systems that influence the global climate. Part of this cycle includes evapotranspiration, which is the net combination of evaporative processes and plant transpiration. The summation of these effects accounts for the atmospheric water vapor content. Globally, 90% of the atmospheric water vapor content is attributed to evaporation from bodies of water (i.e. oceans, lakes, rivers, ect.), and the remaining 10% results from plant transpiration (USGS, 2010). However, over forested regions, a majority of atmospheric water vapor is produced by plant transpiration.

Atmospheric water vapor is important for a variety of reasons. First, it controls precipitation and surface temperature. Second, it is the most abundant greenhouse gas in the atmosphere (NOAA, 2010). It is part of a 'positive feedback loop' that affects the climate system and contributes to global climate change (NOAA, 2010). Combined, these factors could significantly impact the global hydrologic cycle thus influencing global climate change.

Trees are a significant factor in the global hydrologic cycle because they are responsible for transporting water from the soil back into the atmosphere. Indeed, trees can lose up to 95% of their absorbed water through transpiration (Kozlowski et al. 1979). Transpiration is controlled by

a number of physiological factors and environmental factors. These physiological factors include leaf area, root to shoot ratio, leaf size and shape, leaf orientation, leaf surfaces, and stomata, while the environmental factors include light intensity, vapor pressure deficit (VPD) between the stomata cavity and the air outside the leaf, temperature, turbulent mixing of the air around the leaf, and soil water supply (Kozlowski et al. 1979). These factors are coupled with forest heterogeneity and forest canopy structure to help determine the amount of water loss by transpiration a specific type of forest experiences under given meteorological conditions. Changes in a forest's species composition, its canopy height, and the density of the vegetation can affect this amount of water loss (Kozlowski et al. 1979). Predicting how natural forest processes (i.e. species succession, disturbance) and climate change will impact water use and water loss in a forest will be important in predicting the fate of some forests as they succumb to environmental stresses, and for predicting the weather and water supply availability in forested regions.

Recent studies have been conducted to identify how environmental stresses can impact water use and water loss by mature trees. In a study by Bovard et al. (2005) it was shown that daily species-level transpiration (E_i) declined when both photosynthetically active radiation (PAR) and mean daytime (VPD) were low. This means that E_i was positively correlated to the environmental factors PAR and VPD. Soil water content had no effect on E_i at the daily time scale. They also found that sap flux (J_s) increased exponentially with increasing VPD < 1 kPa, and was independent of soil water content. However, they found when VPD > 1 kPa, J_s was sensitive to soil water content. This suggests there is a threshold at which J_s has a more acute response to water availability in the soil, and as VPD increases the water pressure within the

tree's xylem experiences greater stress. Thus, PAR, VPD, and soil water were environmental factors that affected the ecosystem water fluxes (Bovard et al. 2005). Tang et al. (2006) also came to similar conclusions.

In another study by Rocha et al. (2004), cloud cover's effect on photosynthesis (P), evapotranspiration (E), and water use efficiency (WUE) were investigated. They determined that E significantly decreased with increasing cloud cover but P only slightly decreased. As a result, WUE doubled. Moreover, cloud cover helped explain why P was more sensitive or less sensitive to soil water content (Rocha et al. 2004). Thus, if regions become more cloud covered or less cloud covered, these environmental factors will affect water use in trees. If precipitation levels change on a regional scale, soil water content will also change. As more environmental factors interact, the hydraulic cycle will be altered in response to regional and global climate change.

Better understanding of water stress within trees is needed to determine the potential impacts of changes to soil moisture on transpiration and its feedback to the hydrologic cycle. In part, this understanding can come from detailed models of a tree's hydraulic framework. Information about sap flux, transpiration, soil water content, and leaf area are already incorporated in models using empirical equations to predict forest water use. However, these models do not currently incorporate the physical mechanism of water flow in the tree and assume a direct empirical link between leaf responses to soil moisture. The tree's hydrological structure is therefore, not included in these models, and species-specific differences in the allometry of the hydrological system are not included in their calculations of water use. The branching network of a tree can affect the efficiency and maximal rate at which particular species can transport water to its

leaves. When transpiration rates are higher than the rate of water supply to the branch carrying the transpiring leaf, water potential in the leaves and branch will drop and eventually lead to hydrodynamic stress that can damage the tree's conductive system. This damage is caused by intrusion of air bubbles to the conductive tissue, a phenomenon called cavitation. The location of the branches throughout the canopy and how many leaves and leaf bearing branches (LBB) are attached to these branches becomes important in determining the extent of hydrodynamic stresses each tree will experience in response to environmental conditions. Trees have evolved to minimize the risks from hydrodynamic stresses. When pressure in the leaf carrying branch is becoming too low, stomata on the leaves will close, and therefore reduce the demand for transpired water, and allow the pressure to build up again. Because it controls the stomata, the water pressure gradient within the 'pipe network' of the branches and stem is one factor that will control the levels of transpiration (Sperry et al. 2000). Different species have different hydrological structures and different sensitivity to hydrodynamic stress. As forest composition and canopy structure changes, those changes in species and structure can change the water use of the forest. Being able to incorporate species-specific parameters in larger-scale models will improve those models' capabilities to predict transpiration rates. Effects of change to composition of species-specific hydrodynamic structure can be represented by adding or removing specific tree architecture from the model.

Through our research, we predict that a specific tree species will have common allometric relationships even though individuals within a species can be observationally very different. These allometric relationships can then be used to determine a set of parameters that define a specific hydraulic structure. We will use a model that uses these structural parameters and a

random-fractal construction algorithm to generate virtual trees, which represent trees of different species and sizes in our forest plots in UMBS. Another model will use these virtual trees to predict the transpiration rates from our forest. Furthermore, we predict that as a particular type of architecture is introduced to a canopy or succeeds from a canopy, the water use within that forest will change as well.

1.1. Materials and Methods

1.1.1. Study Site

The study sites were two mixed-hardwood forest locations at the University of Michigan Biological Station (UMBS) in Pellston, MI (45°33' N, 84°42' W). Both sites had well-drained spodosols (92.9% sand, 6.5% silt, 0.6% clay) (Bovard et al. 2005). One location was at the AmeriFlux site. The other location was at the Forest Accelerated Succession ExperimenT (FASET) site. At the FASET site all aspen and birch had been stem girdled in a 33 ha plot. At each site three species were measured: Populas grandidentata Michx. (bigtooth aspen); Acer rubrum L. (red maple); and Pinus strobus L. (white pine). Within each species, a range of tree sizes based on diameter at breast height (DBH) were measured.

1.1.2. Measurements

For each tree, multiple measurements were taken in order to characterize its architecture. The measurements were taken utilizing a digital caliper, digital protractor, 50-m measuring tape, 52-

cm measuring tape, and a 71-in collapsing measuring stick. In order to reach all heights along the primary stem, a zero emission canopy access vehicle (ZECAV or “zeke”), Swedish climbing ladders, and tree climbing harnesses were used. The following species-specific tree architectural components were measured and recorded: DBH, crown diameter (CD), primary branch orientation (PBO), primary branch number (PBN), primary branch height (H_B), stem diameter at the height (Z) where a primary branch splits (D_Z), primary branch diameter (D_B), location (length from branch split) of first secondary branches (L_{SB}), primary branch length (L_B), first secondary branch diameter (D_{SB}), primary branch inclination (PBI), total number of secondary branches per primary branch (T_{SB}), and the total number of leaf bearing branches per primary branch (T_{LBB}). In addition to these architectural components, the leaf patterns of each species was also noted (i.e. the number of leaves per leaf bearing branch). For clarification, a primary branch is one that is connected directly to the main stem. A secondary branch is one that is connected directly to a primary branch. The PBO is 1 of 8 cardinal directions to which the primary branch most nearly faces (i.e. north, northeast, east, etc.). These cardinal directions were entered as degree measurements into Excel spreadsheets later in the study (i.e. north=0°, northeast=45°, east=90°, etc.).

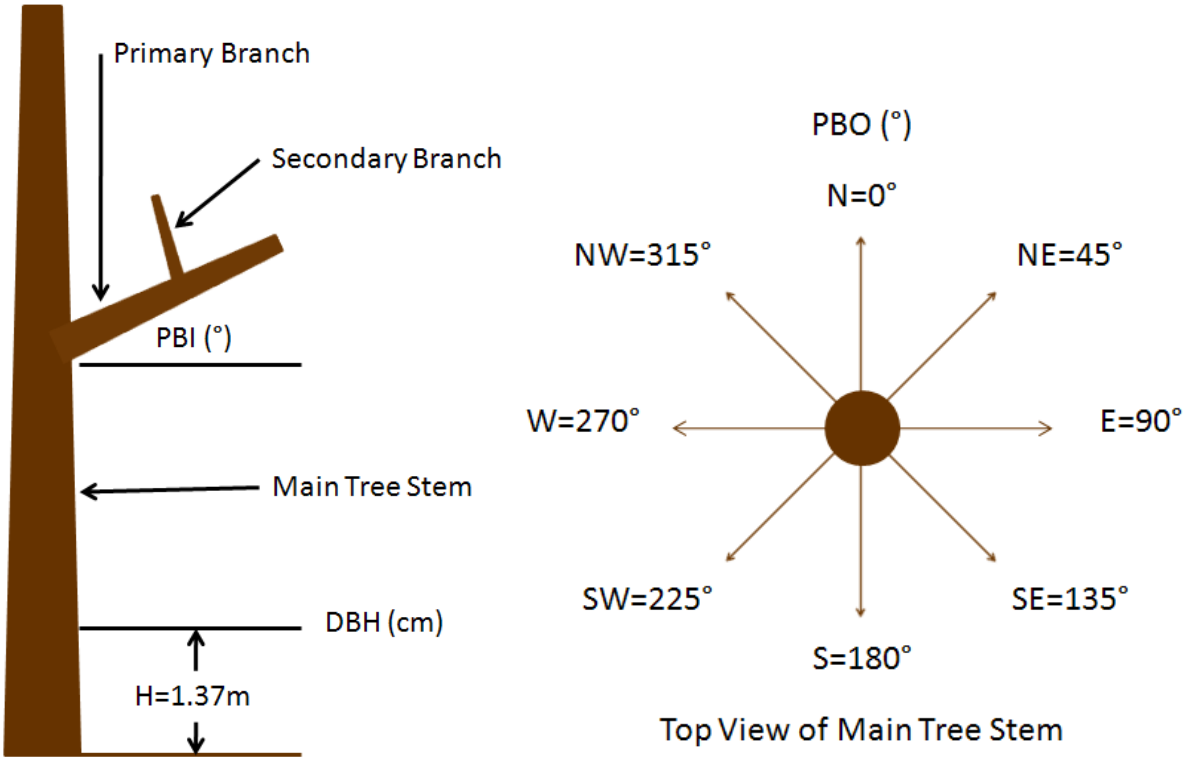

Figure 43: Description of Specific Field Measurements

1.1.3 Methodology

Each individual species' measurements were cataloged into an Excel spreadsheet. Each tree was then used to make comparisons of different measurements in a single tree, between trees of the same species, and between trees of different species. The following comparisons were plotted: individual tree whorl diameter ratios and species-specific whorl diameter ratios (a whorl's total D_B plus D_Z, compared to the preceding whorl's total D_B plus D_Z); species-specific cumulative L_B vs. both DBH and H; species-specific D_B vs. D_{SB}; species-specific D_B vs. T_{SB}; species-specific D_B vs. L_B; species-specific L_B vs. T_{SB}; species-specific H and CD vs. DBH and collective H and CD vs. DBH. Once all plots were created, a MATLAB 'fit' function was used to determine how well a linear or logarithmic function fit the data. Additionally, a tapering function was calculated

for each tree to show the relationship between D_Z and stem heights (H_Z). An existing taper function developed by Ormerod (1973) was used:

$$d(h)^2 = a^2 D^2 \left[\frac{(H-h)}{(H-h_b)}\right]^b \quad \text{Equation (1)}$$

The square root of Equation (1) was taken, thus yielding the following equation:

$$\frac{d(h)}{D} = a\left[\frac{(H-h)}{(H-h_b)}\right]^{b/2} \quad \text{Equation (2)}$$

The left side and right side of Equation (2) were separated into two variables:

$$y = \frac{d(h)}{D} \quad \text{Equation (3)}$$

$$x = \left[\frac{(H-h)}{(H-h_b)}\right] \quad \text{Equation (4)}$$

The 'x' and 'y' functions were plotted against each other in order to calculate the corresponding 'a' and 'b' values for each individual tree. From this information, tree specific taper functions were calculated.

Once all comparisons were made, the linear or logarithmic functions that fit the data were used to determine the strength of allometric relationships between different tree measurements based on the adjusted R-square values generated by MATLAB. Species-specific comparisons included trees of varying DBH, therefore, the identified allometric relationships represented trees of varying sizes.

The identified allometric relationships were then utilized to establish a list of parameters that generally define a particular species with a certain DBH and CD. These species-specific parameters were entered into a MATLAB virtual tree function (Bohrer et al.2005). This function was modified for this project, to allow the generation of highly realistic trees. An Excel

spreadsheet of the branching network for each tree was also created as an input file to aid the MATLAB tree function in developing more species-specific tree structure. This spreadsheet cataloged the branching network of each tree in a manner such that all branches were represented by a line segment with a beginning point and terminal point. Both beginning points and terminal points had a corresponding (x, y, z) coordinate designation. Furthermore, each segment was numbered so that its preceding segment could be numerically identified. A MATLAB loop function within the MATLAB tree function used the Excel spreadsheet files as one of the data sources to help generate a plot of the architectural framework of an individual tree. Finally, basic tree characteristics (i.e. height, average PBI, and mean diameter of stem) were inputted into the virtual tree function. The output was a linear plot representation of a 3-D tree. In summary, the primary goal was to input the characteristics of a particular tree species into the MATLAB tree function and have that function output a numeric array representing a virtual tree structure that physically resembles the desired tree species and a line plot illustrating the virtual tree.

1.2. Results

The significance of the allometric relationships was largely determined by R-square values. The fit functions calculated by MATLAB use a nonlinear least squares method to determine the best fit. The fit functions produce R-square values and adjusted R-square values. It should be noted that the R-square values presented in this analysis are the adjusted R-square values. Many coefficients were added to our function models; therefore, the adjusted R-square statistic was a better indicator of the fit quality (MATLAB, 2010). MATLAB adjusted R-square values can be less than or equal to 1. As with general R-square statistics, the closer the value is to 1 the better

the fit. Any negative adjusted R-square values presented here correspond to extremely poor fits. Table 1 summarizes the DBH, total tree height (H), and CD of all trees analyzed in this study.

The first allometric relationships analyzed were taper functions. The (9) Bigtooth Aspens analyzed had an average ‘a’ value of 1.01 with a standard deviation of 0.03. The average ‘b’ value was 0.85 with a standard deviation of 0.17 (Table 1). The R-square values for the Aspens ranged from 0.4551 to 0.9753 (Figure 2). The (8) Red Maples analyzed had an average ‘a’ value of 0.99 with a standard deviation of 0.08. The average ‘b’ value was 0.94 with a standard deviation of 0.20 (Table 1). The R-square values for the Maples ranged from 0.8605 to 0.9952 (Figure 3). The (8) White Pines analyzed had an average ‘a’ value of 1.01 with a standard deviation of 0.03. The average ‘b’ value was 1.30 with a standard deviation of 0.08 (Table 1). The R-square values for the Pines ranged from 0.9870 to 0.9993 (Figure 4).

The next allometric relationships analyzed were whorl diameter ratios. Each tree’s whorl diameter ratios were plotted and fit to a linear function. The (9) Bigtooth Aspens had R-square values ranging from -0.6220 to 0.7104 (Figure 6). The (8) Red Maples had R-square values ranging from -0.2127 to 0.9171 (Figure 7). The (8) White Pines had R-square values ranging from 0.3328 to 0.9083 (Figure 8). Each tree was also combined into a collective data set representing the overall whorl diameter ratios for each species (Figure 9). The Aspens collectively had 106 whorls, with an R-square value of -0.1629. The Maples collectively had 168 whorls, with an R-square value of 0.9761. Finally, the Pines collectively had 217 whorls, with an R-square value of 0.9809.

The remaining allometric results presented here are comparisons between different types of tree measurements. For all of these comparisons the sample size will be significantly different between the three species. It should be noted that from (9) Bigtooth Aspens, 98 primary branches were measured. From (8) Red Maples, 232 primary branches were measured. From (8) White Pines, 703 primary branches were measured.

First, D_B was compared to D_{SB} (Figure 10). All three species were fit to a linear function. The Bigtooth Aspen's R-square value was 0.6047. The Red Maple's R-square value was 0.9887, and the White Pine's R-square value was 0.9916. Next, D_B was compared to T_{SB} (Figure 11). All three species were fit to a linear function. The Bigtooth Aspen's R-square value was 0.9790, the Red Maple's R-square value was 0.9878, and the White Pine's R-square value was 0.9927. Then, D_B was compared to L_B (Figure 12). All three species were fit to a linear function. The Bigtooth Aspens' R-square value was 0.7707, the Red Maple's R-square value was 0.9946, and the White Pine's R-square value was 0.9974. Finally, each species' L_B was compared to T_{SB} (Figure 13). The Bigtooth Aspen's R-square value was 0.4557, the Red Maple's R-square value was 0.9904, and the White Pine's R-square value was 0.9987.

The PBI and PBO for each species were analyzed differently and did not utilize MATLAB fit functions. Instead, basic statistical test were used for both inclinations (Figure 14) and orientations (Figure 15). The Bigtooth Aspen's average PBI was 41.23° with a standard deviation of 16.53°. The Red Maple's average PBI was 26.76° with a standard deviation of 26.77°. The White Pine's average PBI was 11.76° with a standard deviation of 20.37°. The Bigtooth Aspen's PBO was 155° with a standard deviation of 102°. The Red Maple's PBO was

159° with a standard deviation of 97°. The White Pine's average PBO was 161° with a standard deviation of 104°.

The next analysis involved cumulative L_B compared to both DBH and H (Figure 16). All data points were fit to a linear function. The Bigtooth Aspens had an R-square value of -0.8344 for its DBH comparison and a value of -0.2573 for its H comparison. The Red Maples had an R-square value of 0.5715 for its DBH comparison and a value of -0.2486 for its H comparison. The White Pines had an R-square value of 0.8330 for its DBH comparison and a value of 0.9118 for its H comparison.

The last allometric relationship studied involved the comparison of both H and CD to DBH. Plots were created specific to tree species (Figure 17) and inclusive of all tree species (Figure 18). This analysis incorporated 35 Bigtooth Aspens, 38 Red Maples, and 39 White Pines. All curves fit to these datasets were logarithmic functions. The Bigtooth Aspens, collectively, had an R-square value of 0.9281 for the H comparison and 0.3803 for the CD comparison. The Red Maples, collectively, had an R-square value of 0.5044 for the H comparison and 0.2163 for the CD comparison, The White Pines, collectively, had an R-square value of 0.8747 for the H comparison and 0.9621 for the CD comparison. When all trees were combined, the R-square value was 0.9927 for the H comparison and 0.9803 for the CD comparison.

After the allometric relationship comparisons were completed, all of the data was combined into a finite element table. This table cataloged every primary branch (element) and stem portion (i.e. stem length between whorls) as a line segment with starting and ending (x, y, z) coordinates. A

MATLAB function read this data and generated 3-D plots of the different tree species. These plots can be viewed in Figures 19, 20, and 21.

1.3. Discussion and Conclusion

After analyzing the results, we found there are many allometric relationships that exist within an individual tree, within a tree species, and between tree species. Some of these relationships had stronger correlations than others. Taper functions were important to generate because they help predict how D_Z changes with H_Z. Looking at the individual tree taper functions and the summary plot (Figure 5), we saw distinct differences between species. Aspens were more likely to fit a general taper function representing all individuals within the sample based on their 'a' and 'b' values. The 'a' and 'b' values of the Maples, however, show that it would be difficult to represent all individuals with one general taper function. Compared to Aspens and Maples, the 8 Pines can be individually identified in the summary plot (Figure 5). Moreover, Pines had the smallest standard deviation in 'b' values which supports stronger uniformity among Pines. This analysis shows that Aspens and Pines are more likely to fit general taper functions, whereas Maples do not fit general taper functions well.

Whorl diameter ratios were calculated to see how the cross sectional areas of one whorl would change in proportion to the cross sectional areas of the next whorl along the tree stem. Using a circle as area representation, diameter ratios would be directly proportional to cross sectional area ratios. On an individual tree basis, there was great variability in how well a linear function fit the ratios. Of the 3 species, White Pine had the only consistent R-square values that would

support a linear relationship. Aspens and Maples showed a higher level of variability. When the trees of each species were combined, the results were different. Aspens still had no significant correlation between one whorl and its preceding whorl. Based on R-square values, however, both Maples and Pines showed a strong linear relationship among their ratios. As a population, one could estimate how the cross sectional areas at one height would be related to that of another height.

The remaining allometric relationships had more significant results. We expected there to be strong correlations with D_B compared to D_{SB}, T_{SB}, and L_B. This was confirmed by the R-square values for all linear fits to each analysis. For all species there was a positive correlation for all comparisons. As D_B increases, D_{SB} increases, T_{SB} increases, and L_B increases. Furthermore, as L_B increases, T_{SB} increases. With regard to PBI and PBO, we did not expect significant results. We did, however, predict different average PBI for each species. This prediction was confirmed by the statistical tests. There was a difference in average PBI between the species where Aspen had the largest PBI and Pine had the smallest PBI. However, the standard deviations were so great for each average PBI that the results were not conclusive. The PBO analysis showed all species having a similar average PBO, but the standard deviations were large enough to prevent a correlation from being significant.

The analysis comparing cumulative L_B to both DBH and H helped categorize the 3 species. It should be noted that for each branch an error of ±0.45m was incorporated into the cumulative calculation. This, however, only resulted in a maximum of ~2% error in the cumulative calculation. For the Aspen, the R-square values support that one cannot estimate cumulative L_B

based on the tree's DBH or H. For the Maple, R-square values support that one can estimate cumulative L_B based only on the tree's DBH. One cannot estimate a Maple's L_B based on the tree's H. For the Pine, the R-square values support that one can estimate cumulative L_B based on the tree's DBH or H. These results support that more uniformity exists in White Pines with respect to cumulative L_B.

The final allometric analysis comparing H and CD to DBH helps support that relationships exist within species and between species. The comparisons were fit to logarithmic functions and the corresponding R-square values are conclusive. For the 3 species, Aspens and Maples had stronger correlations between H and DBH, whereas Pines had stronger correlations between CD and DBH. Regardless, all comparisons made within species had strong relationships supported by high R-square values. When all species were combined, the results were still significant. The R-square values strongly support that logarithmic relationships exist between H, CD and DBH.

The 3-D 'pipe models' created helped us visualize the architecture of a specific tree species. The allometric relationships developed beforehand couple well with the produced images. One can see that of the 3 species, White Pine is the most uniform. Aspen have fairly uniform structure, but vary in branch arrangements at the top of the tree, while Maples show the greatest amount of variability in their architecture. These visual results can be further validated by the support shown in the allometric analyses.

These 'pipe network' models of trees could be used to start enhancing models of the hydrological cycle. The models show that hydrodynamic structures of trees vary between tree

species. This hydrodynamic structure could have an integral role in determining how rates of transpiration change in response to water stresses experienced within the tree's 'pipe network'. But before it can be supported that the 'pipe network' of a tree has a specific influence on the water stresses a tree experiences, environmental factors must also be considered. This 'pipe network' model should be coupled with sap flux measurement, soil moisture content, transpiration rates from leaves, and turbulent mixing of air around the leaves. Moreover, the tree's anatomy and physiology must be considered because different tree species have different types of tissue affecting the porosity within the 'pipes'. This would help define what type of 'pipe' a certain tree species has. When all of these measurements are combined a more complex model of transpiration responding to water stresses could be developed. We need to find out what the most limiting factor to transpiration is in response to water stress. If we consider the leaves as 'valves', the branches as 'pipes', and the soil below as a 'reservoir', then this future work could determine which 'structure' is the limiting factor. These findings could greatly enhance the predictive capabilities of hydrological cycle models, which would be crucial as climate changes and forest composition changes.

Chapter 2: Tree-type classification of high-resolution airborne images combining visible and LIDAR

2.0. Introduction

Allometric relationships are one step to developing more robust input parameters to existing biosphere-atmosphere models. New generation models will require such inputs to enhance their predictive capabilities of the exchanges that occur between forested ecosystems and the atmosphere above. Species-specific structure and physiology will add to the depth of future input parameters making a modeled forest more real where it is based upon both empirically and observationally derived data. But these input parameters at an individual-tree scale require census information as well. Once you create computer-generated trees that display structural and physiological characteristics derived from field measurements, you need to know how to better classify a computer-generated forest. The next step is utilizing remote-sensing data to classify a forested region to some level of confidence based upon the canopy dominant species of that region. When a model can have a realistic individual tree and its associated characteristics in addition to how many of those trees exist in a specified region, the model becomes a powerful tool. This tool can work with real field data and empirically derived relationships which strengthen the validity of its results. As such the second part of this research determined whether or not remote-sensing is a viable option for classifying forested regions with variable heterogeneity.

2.0. Materials and Methods

2.0.1. Software

The remote-sensing analysis for this research can be attributed to the computational image processing capabilities of ENVI (version 4.5). ENVI has a unique set of image analysis tools that enable the user to manipulate a generic image input file and output visual results backed by statistic packages that focus on specific aspects of each unique image input file. This research project utilized the following ENVI tools:

- Image Resizing
- Region of Interest (ROI)
- Texture Filter
- Mosaic
- Layer Stacking
- Masking
- Classification
- Post-Classification

Details on each of these tools will be described in the methodology and result interpretation subsections.

2.0.2. Methodology

The first step toward remote-sensing analysis utilizing aerial photography was selecting an appropriate image set with which to work. The primary factor taken into consideration at this step was a cost-benefit comparison between the actual cost of an aerial image dataset and its corresponding resolution. Though higher resolution images would afford the user more confidence in identifying specific features in each image, the cost of such resolution may not justify the use of such images. After careful consideration, it was agreed that imagery provided by both a cost-free and publically accessible image database was most appropriate. We elected to use the National Agriculture's Imagery Program (NAIP) database due to its ease of accessibility

and breadth of land surface coverage. NAIP imagery datasets can be selected based upon the United States County of interest, as such, datasets for years 2010, 2009, and 2006 of Cheboygan County Michigan were downloaded. Furthermore, each image dataset needed specificity with regard to its coordinate setup system. For the purpose of this research we chose to work in Universal Transverse Mercator (UTM) coordinates relative to the North American Datum of 1983 (NAD 83) Zone 16 where Zone 16 corresponds to the region of Michigan in which UMBS is located. Both the 2010 and 2009 datasets had pixel resolutions of one meter whereas the 2006 dataset had a two meter pixel resolution.

Once our image datasets had been selected we used ENVI to resize the image to a region only covering our research site. This is accomplished by establishing a Region of Interest (ROI) in ENVI. We selected the region encompassing the AmeriFlux research site at UMBS and saved that region as a specific ROI. Resizing the image to a smaller coverage area subset greatly reduced the number of pixels associated with each image dataset. This was important from an efficiency standpoint because when we ran classifications of image regions further in the research, the reduced number of pixels permitted ENVI's analysis to take less time than would be anticipated had the entirety of Cheboygan County been analyzed.

Figure 44: 2010 NAIP Image with Douglas Lake Masked

For computational processing of image files, every one-meter-by-one-meter pixel in an image is represented by three color bands: red, green, and blue (RGB). Each pixel, therefore, has one numerical value for each color band representing that pixel. RGB colors can have one of 256 distinct values ranging from 0-255. For a pixel to be visually perceived as black, all RGB values in that pixel must be 0. For a pixel to be visually perceived as white, all RGB values must be 255. Because every pixel is unique, every image inherently has textures associated with it which represent the statistical differences between adjacent pixels. As such, each texture can be filtered out and used independently of the other textures. This is accomplished by using another tool in ENVI that filters the textures of an image into separate components. When the components are separated, fifteen bands will be produced, five for each RGB color band. These textures include data range, mean, variance, entropy, and skewness. For image processing, we elected to utilize variance bands for each color band. The variance bands proved to be a useful input parameter in our imagery analysis because the variance bands help distinguish the visible gaps in the forest

canopy when viewing the forest from an aerial perspective. In essence, the RGB variance bands helped identify the change in canopy structure with regard to visual color change.

With RGB bands and variance bands for those colors from three image datasets representing different NAIP imagery years, we still wanted to add more bands before we began classifying our research site. Based on our allometric results, the most logical input would be aerial Laser Imaging Detection and Ranging (LIDAR) data because the height returns could be correlated to our allometric relationships found between DBH, tree height, and crown diameter. The aerial LIDAR data covered a region with a width of ten kilometers and a height of seven kilometers. For ease of inputting this data in to ENVI, we divided this large region in to equally sized subregions with widths of two kilometers and heights of one kilometer. By georeferencing a research site with the LIDAR data we only had to input nine of the thirty-five created subregions. Once all nine subregions were inputted to ENVI, we used ENVI's mosaic tool to merge these subregions in to one master region. This LIDAR region became the last band we used as an input before image classification.

In total, we used thirteen unique bands for our image classification which are as follows:

1. 2006 NAIP Red Color Band
2. 2006 NAIP Green Color Band
3. 2006 NAIP Blue Color Band
4. 2009 NAIP Red Color Band
5. 2009 NAIP Green Color Band
6. 2009 NAIP Blue Color Band
7. 2010 NAIP Red Color Variance Band
8. 2010 NAIP Green Color Variance Band
9. 2010 NAIP Blue Color Variance Band
10. 2009 Aerial LIDAR Mosaic Band
11. 2010 NAIP Red Color Band
12. 2010 NAIP Green Color Band
13. 2010 NAIP Blue Color Band

Figure 45: 2006 NAIP RGB Layer Reduced to Research Site

Figure 46: 2009 NAIP RGB Layer Reduced to Research Site

Figure 47: 2010 NAIP RGB Layer Reduced to Research Site

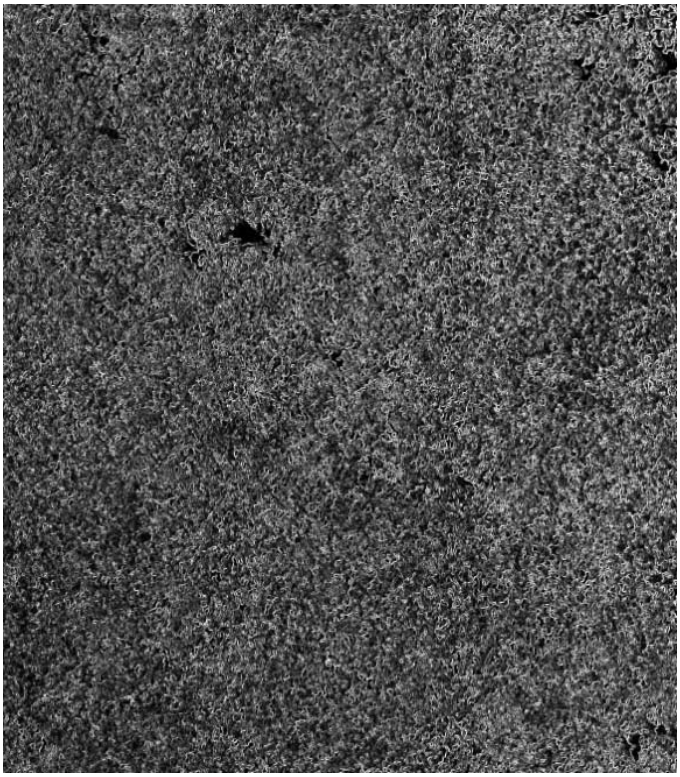

Figure 48: 2010 NAIP Red Variance Layer Reduced to Research Site

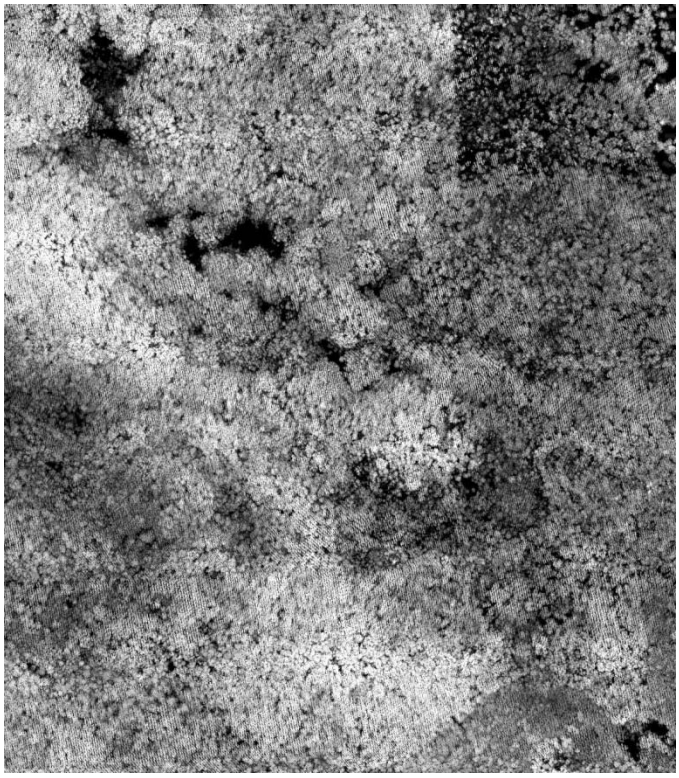

Figure 49: 2009 Aerial LIDAR Data Reduced to Research Site

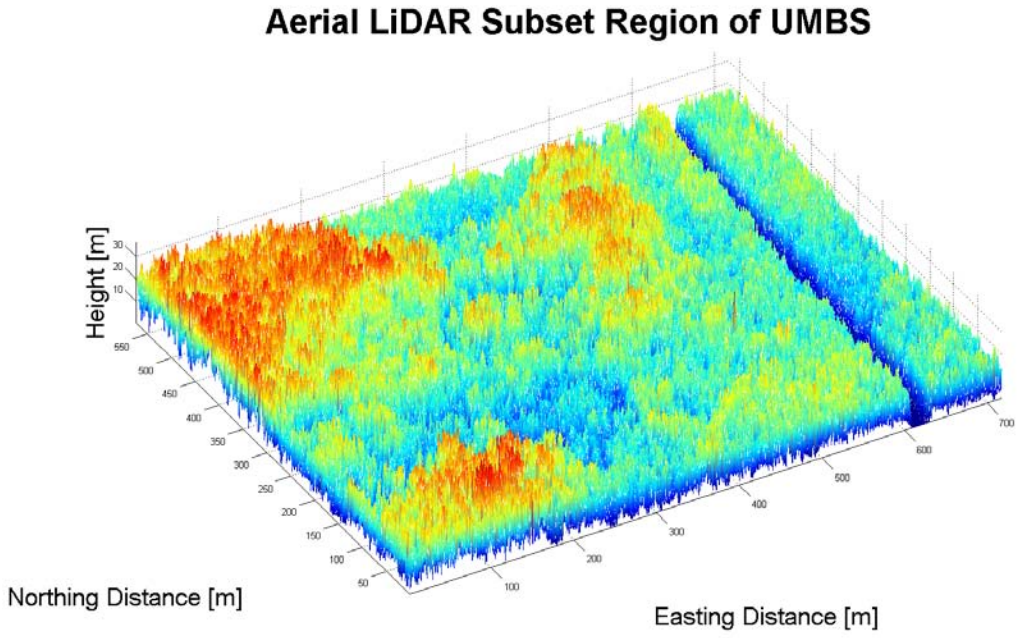

Figure 50: 3D Representation of LIDAR Data Showing Height Variation of UMBS at AmeriFlux

To make a more robust package for classifying our research site, we chose to layer all thirteen bands. ENVI's layering tool takes individual bands and stacks them on top of each other. The end result is an image representing a specific region that has multiple layers within it. By stacking layers, we gave each pixel in the image more depth. The more depth the pixel had the more unique it would be. More simply, instead of having one pixel defined by a RGB band from one year, i.e. 2010 NAIP, we now had one pixel defined by three groups of RGB bands from different years in addition to variance bands and the LIDAR band. Therefore, in classification, it would be less likely that one pixel and an adjacent pixel would have the same parameters in all thirteen bands. If such a similarity did exist, however, we would be more confident that those two pixels represent the same type of identifiable feature in an aerial-based image.

Now that we had an image dataset prepared with thirteen bands we had to start establishing regions of interest that we did not want analyzed. Our research was interested in classifying the forest with remote-sensing data, however, inherent within remote-sensing images are a multitude of other non-forest related features i.e. roads, grass, bodies of water, open areas, buildings, shadow, etc. We created an ROI for Douglas Lake, the primary body of water on the northern boundary of the research station. With this ROI, we built a mask using a tool in ENVI and applied the mask to our layered image file. With a mask in place, all pixel values within the masked region our voided out. What this means, is when we classify the image those masked pixels are not considered as input parameters anymore because they no longer had data associated with any of the thirteen layered bands. This was important in our classification of the region because water, visually, has a lot of variability that could be seen over forested regions as well. If the water was included as an input parameter, more confusion would inherent to our

classification of the forest which decreases the validity of our results. For the purpose of UMBS, we only created three additional ROI's representing regions we did not want to be considered in the forest classification scheme. These regions included road, low-growth (dirt, grass, field, etc.), and shadow. Instead of building a mask for these ROI's, however, we elected to include them within the classification so that shadow within the research plot could be identified.

2.0.3. ENVI Classification and Post-Classification

After creating an image dataset with thirteen band layers and the appropriate ROI's to distinguish forest features from non-forest features, we began looking at the 2010 Census data corresponding to our research site. At the AmeriFlux tower, we used census data from sixty, sixteen meter radius plots located on transects extending at predetermined azimuths from the center of the tower in addition to the sixty meter radius plot surrounding the tower. The next challenge, however, was creating an ENVI-compatible input file that would display census data such as the type of tree and its corresponding DBH. Each plot center had a UTM coordinate set associated with it so we were able to write a MATLAB program that would place all sixty-one plots in spatially correct relationship to one another. In essence, a square matrix was created with a one-meter resolution that displayed the location of every tree in every plot with its species and DBH designation. This file was then used as an additional input to ENVI with geo-referenced coordinates. This means we could simultaneously open our layered image file and our census data file, and geographically link them. With a geographic link in place and a pixel locator tool we were able to select a tree in a specific plot from the census and locate it on the stacked layer dataset of NAIP and LIDAR imagery.

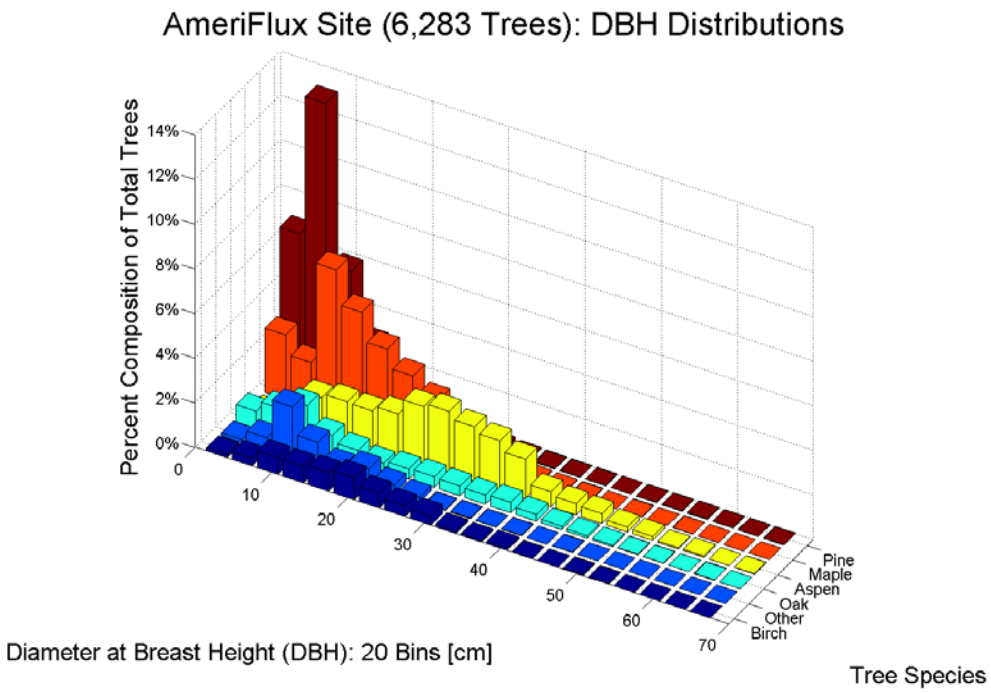

Figure 51: 2010 Census DBH Distribution by Species

The census data was used to establish 'ground truth' data to be later used in testing the validity of our classification results. Looking at the 2010 Census, we decided to create three ROI's representing three different vegetative classes: Aspen, Aspen/Mixed, and Hardwood. Aspen classes represented plots where seventy-five percent or more of the trees in that plot were comprised of Aspen. Aspen/Mixed classes represented plots where less than seventy-five percent but greater than fifty percent of the trees in that plot were comprised as Aspen. Finally, Hardwood classes represented plots where less than fifty percent of the trees in that plot were comprised of Aspen. These ROI's were saved to be used in running a classification analysis. For clarity, every pixel selected within one of these class ROIs added to the definition of that class. These pixels and their representative features from the thirteen bands within them would add depth to the class and hopefully separate them from the other classes.

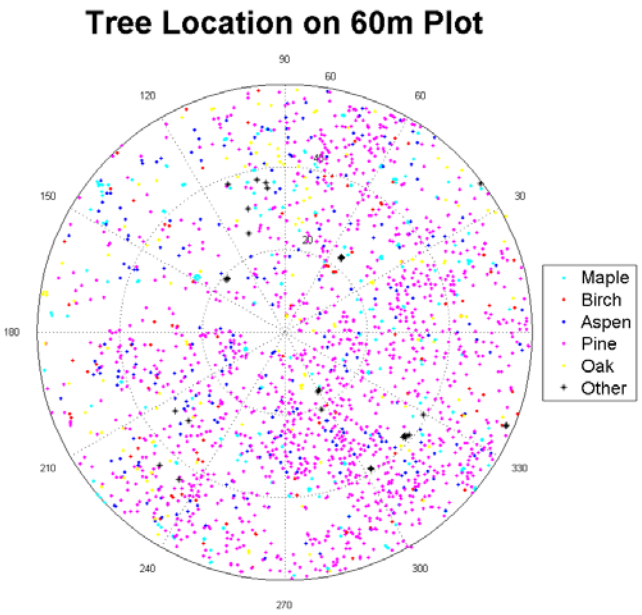

Figure 52: Example of Circular Plot with Tree Species and Location Indicated

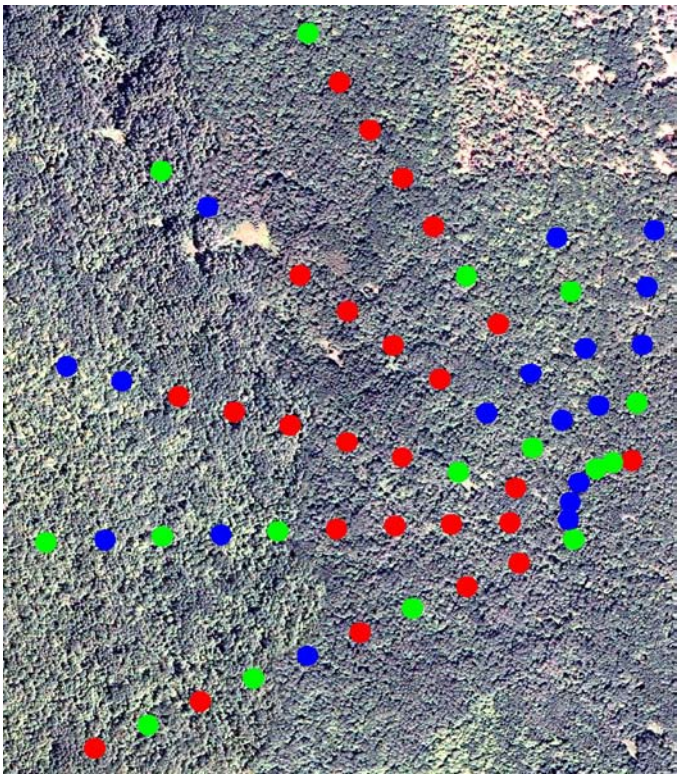

Figure 53: 2009 NAIP Image with 16m Plots Overlayed and Classified as ROI's

To further increase classification efficiency, we resized our layered image to exclude all roads, and building features present in the original image. This area still encompassed all plots within the AmeriFlux research site. To prevent confusion when running the classifications, we elected to include the shadow ROI as well as the low-growth ROI. In ENVI's classification tools, we ran a maximum likelihood classification of the resized image. What this classification accomplishes is using the input ROI's with their corresponding pixel characteristics, and defining the rest of the region based upon pixel similarity. The finished product is an image with every pixel colored to represent the class to which it most nearly belongs based on the ROI input parameters. With our research site classified, we utilized ENVI's post-classification tools to verify the validity of our results. ENVI post-classification tools employ statistical algorithms to compare image input and image output parameters. The primary tool of interest to this research was the post-classification ground truth confusion matrix. This analysis tool uses ground truth data, i.e. the ROI input parameters, and compares those input parameters to how the image was actually classified. Ground truth data can be seen as telling an image what should be on the ground. When a ground truth confusion matrix is run, it reports how often what we said should be on the ground was actually classified as such. For example, if we define 1000 pixels as a single ROI, run a classification, and the confusion matrix says only 800/1000 pixels in that ROI were classified that way, then we were only eighty percent accurate in defining that particular region as user. User accuracy and producer accuracy are an important measure to note in addition to the overall accuracy and kappa coefficient associated with ground truth statistics result. This will be discussed further in the results section.

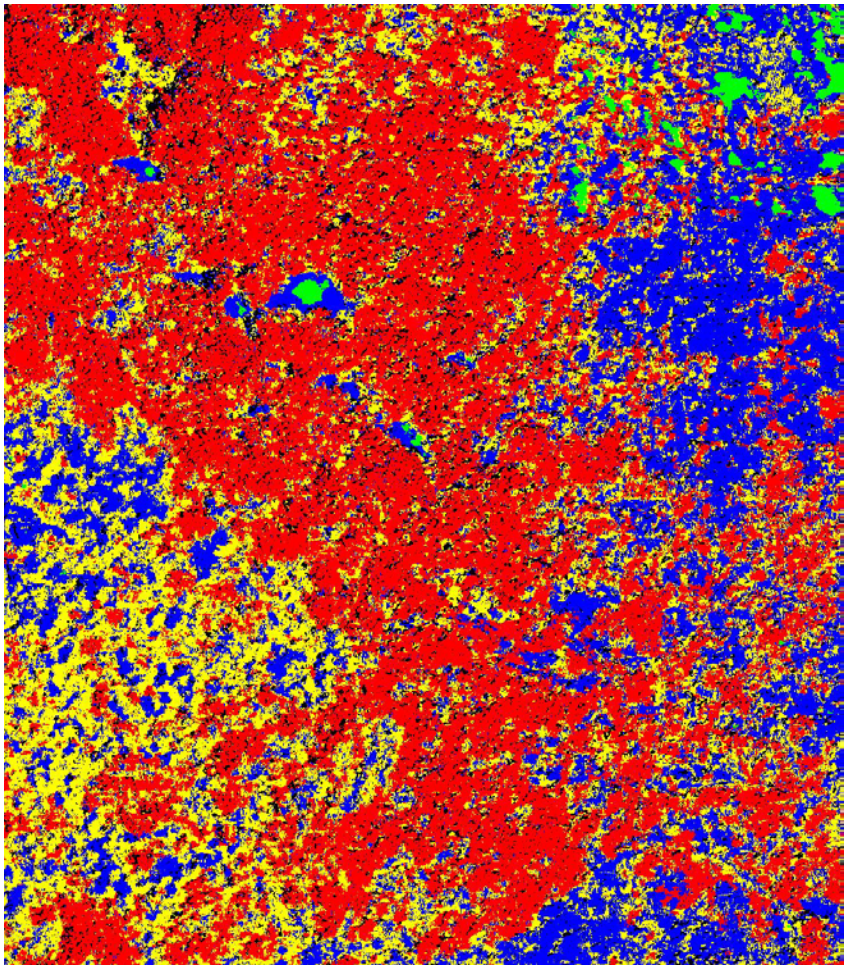

Figure 54: Maximum Likelihood Classification Using 13 Band Layers

Classification Key	
Color	*Description*
Red	Aspen Dominated
Yellow	Aspen/Mixed Dominated
Blue	Hardwood Dominated
Green	Low Growth Areas
Black	Shadow

2.1. Results

Most of the remote-sensing analysis involved setting up the image files with the desired layers and geo-referencing the external data such as the 2010 Census. The next step was running

maximum likelihood classifications with different spectral subset combinations. Though the image contained thirteen bands, the user could elect to use only specific layers when running the classification, i.e. only the 2010 NAIP RGB and LIDAR. Furthermore, we decided to make the shadow ROI more robust by creating it using ENVI instead of the subjective visual acuity of the user. This involved ENVI's band threshold feature to select all red color pixels ranging in value from zero to twenty. These values corresponded to dark shadow regions seen on the image. Through trial and error we ran multiple different combinations of bands and looked at the corresponding post-classification results from the ground truth confusion matrix. In the end, we found that using all thirteen bands produced the best results. The 'best' results were picked according to the combination of the overall accuracy and its subsequent kappa coefficient. The numbers behind the post classification are illustrated in the table below:

	Ground Truth (Pixels)						Producer Accuracy	User Accuracy
Class	*Aspen*	*Aspen/Mixed*	*Hardwood*	*Low Growth*	*Shadow*	*Total*	%	%
Aspen	17588	4956	3705	0	267	26516	68.09	66.33
Aspen/Mixed	3751	6749	3265	0	97	13862	40.56	48.69
Hardwood	2402	3399	10286	5	512	16604	55.28	61.95
Low Growth	3	3	12	3081	1	3100	99.84	99.39
Shadow	2086	1533	1340	0	120480	125439	99.28	96.05
Total	25830	16640	18608	3086	121357	185521		
					OVERALL ACCURACY			**85.26**
					KAPPA COEFFICIENT			**0.72**

Table 2: Post-Classification Ground Truth Confusion Matrix Results from ENVI Processing

To interpret the results we had to look at each ROI class individually as well as the overall summary. We (the user) told ENVI that 26,516 pixels represent Aspen dominated plots and ENVI (the producer) told us that 17,588/26,516 of the pixels were actually classified as Aspen.

This results in 66.33 percent user accuracy. The user accuracy was markedly less with the Aspen/Mixed classification, where we only classified the ground correctly 48.69 percent of the time. This makes sense, however, because we saw that the other half (51.31 percent) of the pixels were distributed rather equally across both the Aspen and Hardwood dominated classes. This led us to see that more specific classes could be used to decrease the amount of confusion, i.e. instead of Aspen/Mixed, have an Aspen/Maple and Aspen/Oak class. The user accuracy for Hardwood dominated plots (61.95 percent) was similar to that of Aspen which was an expected result. The user accuracies for both Low Growth and Shadow were extremely high, which helps validate that our trees were not classified as something other than a tree. In summary, the post-classification told us that we had an 85.26 percent likelihood of classifying the entire image region correctly based on our input ROI's. This accuracy was backed by a kappa coefficient of 0.72.

The ground truth confusion matrix post-classification also helped us determine which bands produced the most separability between the five different input ROI's. The figures below help explain this separability:

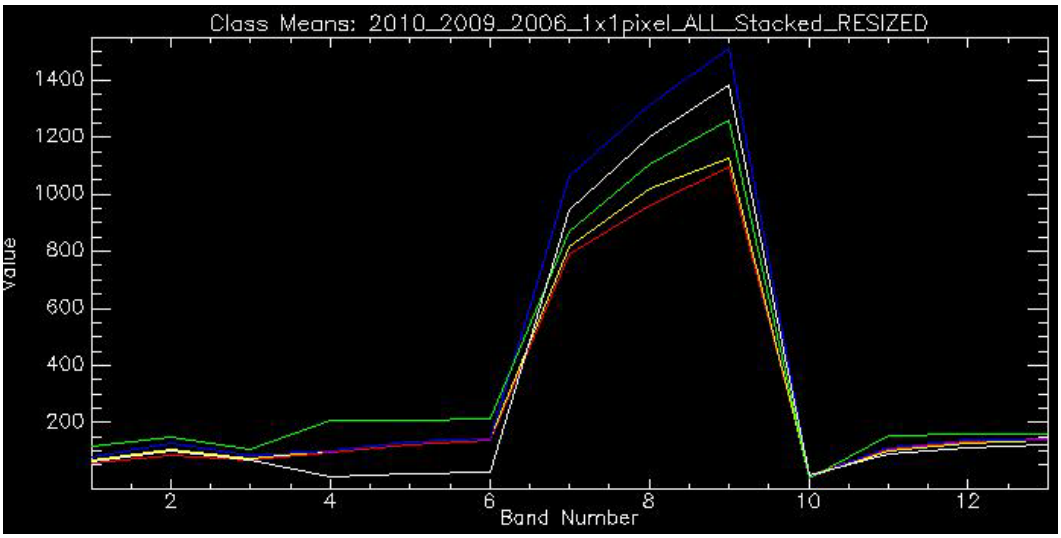

Figure 55: Ground Truth Confusion Matrix Plot Showing Band Separability of the Means

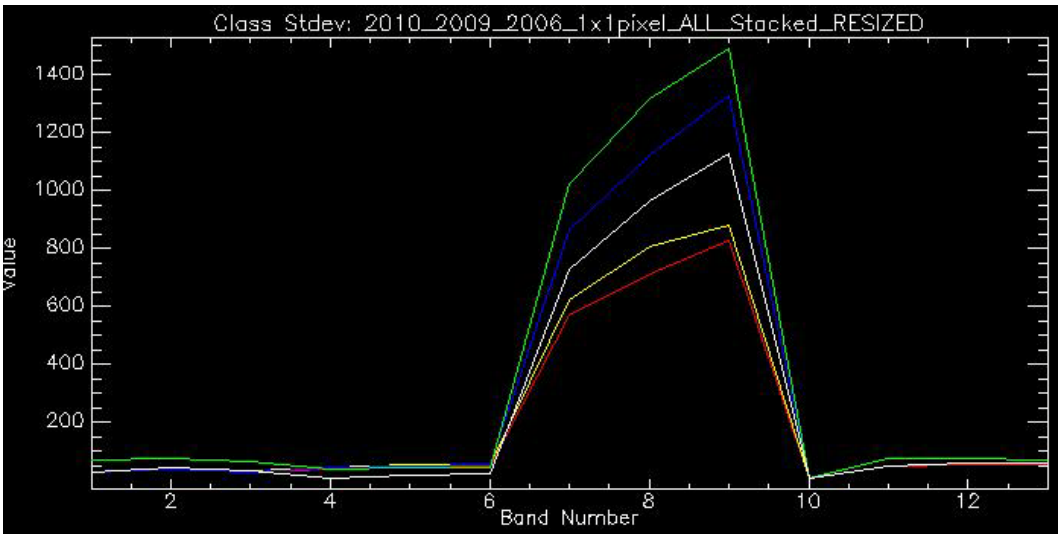

Figure 56: Ground Truth Confusion Matrix Plot Showing Band Separability of the Standard Deviations

When considering both the mean and standard deviation in the separability of the classes, it was clear that band numbers seven, eight, and nine produced the greatest differences between classes. These bands corresponded to the RGB Variance layers in the layered image file. The means figure also depicted that the 2009 NAIP imagery had better separability compared to the datasets from 2006 and 2010. This was subjectively confirmed by visually inspecting the different images

for each year and seeing that 2009 appeared to have better clarity and differentiable features than the images in comparison.

2.2. Discussion and Conclusion

The results from our remote-sensing analysis showed that to some level of confidence, a forested region can be classified based upon a predetermined sample of ground truth data. These classifications, in turn, can be used to both generally define and represent the heterogeneity of a forested region with regard to its dominant species composition. This is useful for future research because existing atmosphere-biosphere models could use such information as more robust input parameters, i.e. giving the models a more true representation of forest composition reflecting individual species versus categorizing a forest as a 'big leaf' in some instances. It should be noted, however, that imagery analysis can be quite subjective. With regard to defining ROI's or eliminating areas of noninterest and other 'noise', what one user may visually see in an aerial image can be different from another user looking at the same image. Therefore, image details that the human eye cannot detect must rely on the processing capabilities of software packages such as ENVI.

Another problem with regard to aerial image analysis is in the nature of the image only showing canopy dominant species. Using imagery such as NAIP to classify a region of interest may neglect the heterogeneity expressed in the understory species. If this data is to be used, the assumption must be clear that canopy dominant species are more crucial in the atmosphere-biosphere exchanges. However, as forests experience natural succession, the canopy dominant

species will change over time. Relating this information back to the allometric relationships that govern a particular species could be a powerful tool in predicting how the ‘new’ forest’s canopy dominant composition could affect exchanges at the boundary layer. Such information will be useful in predicting climate changes at a micro level over areas that experience land use changes, which could potentially impact global scale climate change.

3.0 References

Bohrer, G.: FETCH 2005

Bovard, B. D., P. S. Curtis, C. S. Vogel, H. –B. Su, and H. P. Schmid. "Environmental controls on sap flow in a northern hardwood forest." Tree Physiology 25.1 (2005): 31-38. Electronic.

Kozlowski, T. T., and S. G. Pallardy. Physiology of Woody Plants: Second Edition. San Diego: Academic Press, Inc., 1979. Print.

MATLAB: Documentation: Curve Fitting Toolbox: Evaluating the Goodness of Fit. 19 August 2010. <http://www.mathworks.com/access/helpdesk_r13/help/toolbox/curvefit/ch_fitt9.html>.

National Oceanic and Atmospheric Administration: National Climate Data Center. Greenhouse Gases: Frequently Asked Questions. Water Vapor. 10 July 2010. <http://www.ncdc.noaa.gov/oa/climate/gases.html#watervapor>.

Ormerod, D.W. "A simple bole model." Forestry Chronicle. 49 (1973): 136-138. Electronic.

Rocha, A. V., H. –B. Su, C. S. Vogel, H. P. Schmid, and P. S. Curtis. "Photosynthetic and Water Use Efficiency Responses to Diffuse Radiation by an Aspen-Dominated Northern Hardwood Forest." Forest Science 50.6 (2004): 793-801. Electronic.

Sperry, J.S. "Hydraulic constraints on plant gas exchange." Agricultural and Forest Meteorology. 104 (2000): 13-23.

Tang, J., P. V. Bolstad, B. E. Ewers, A. R. Desai, K. J. Davis, and E. V. Carey. "Sap flux—upscaled canopy transpiration, stomatal conductance, and water use efficiency in an old

growth forest in the Great Lakes region of the United States." Journal of Geophysical Research-Biogeosciences 111.G02009 (2006): 1-12. Electronic.

United States Geological Survey: Water Science for Schools. The Water Cycle: Evapotranspiration. 10 July 2010. <http://ga.water.usgs.gov/edu/watercycleevapotranspiration.html>.

4.0 Appendix A: Figures and Tables

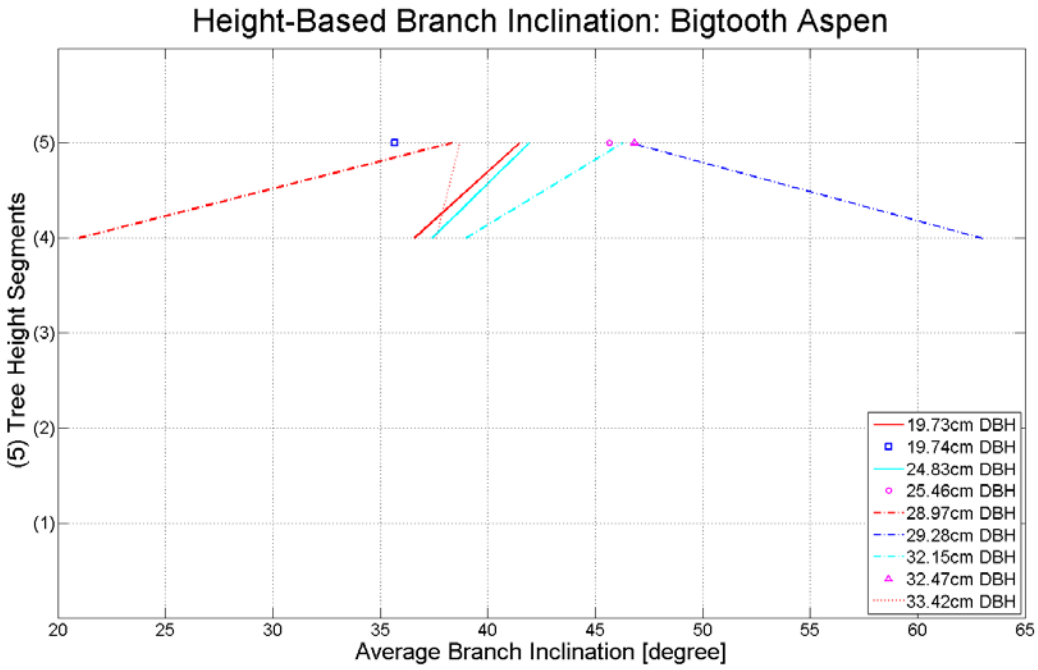

Figure 57: Height-Based Branch Inclination for Bigtooth Aspen by Segment

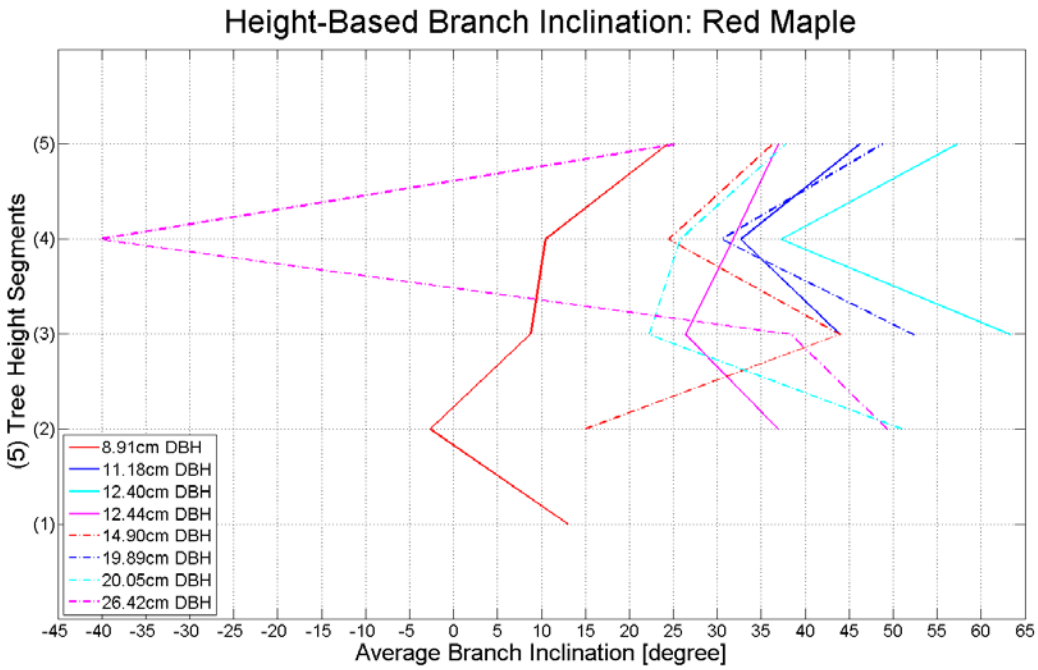

Figure 58: Height-Based Branch Inclination for Red Maple by Segment

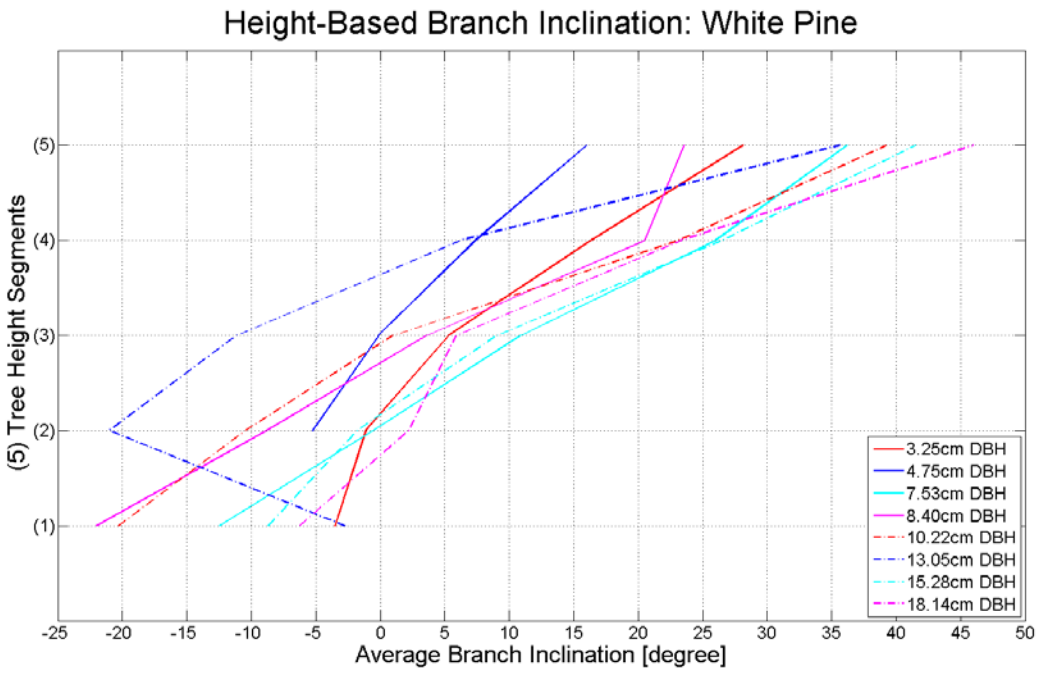

Figure 59: Height-Based Branch Inclination for White Pine by Segment

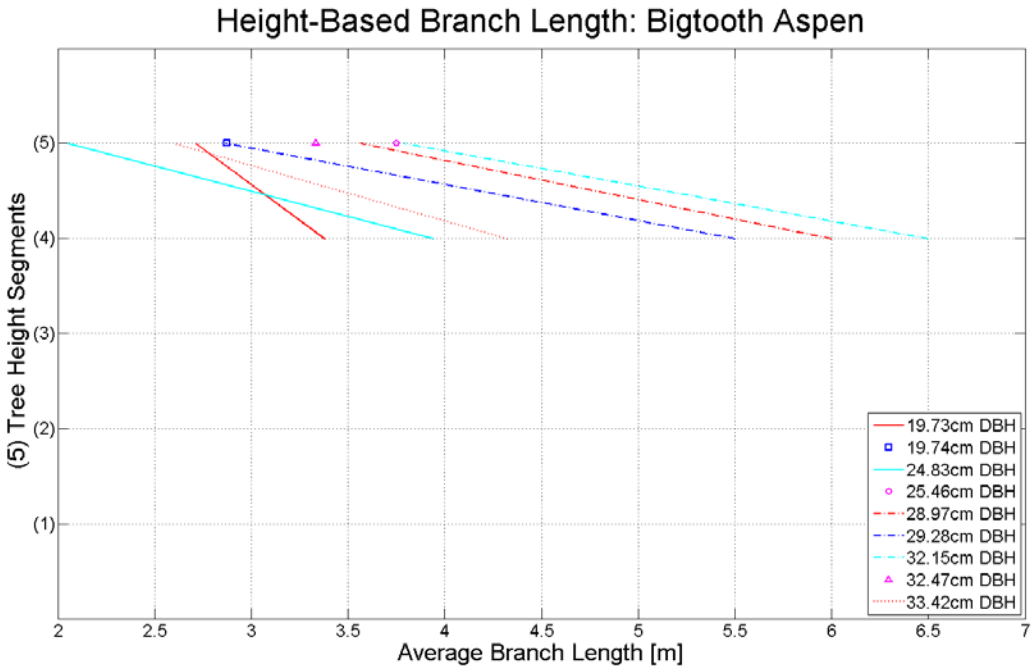

Figure 60: Height-Based Branch Length for Bigtooth Aspen by Segment

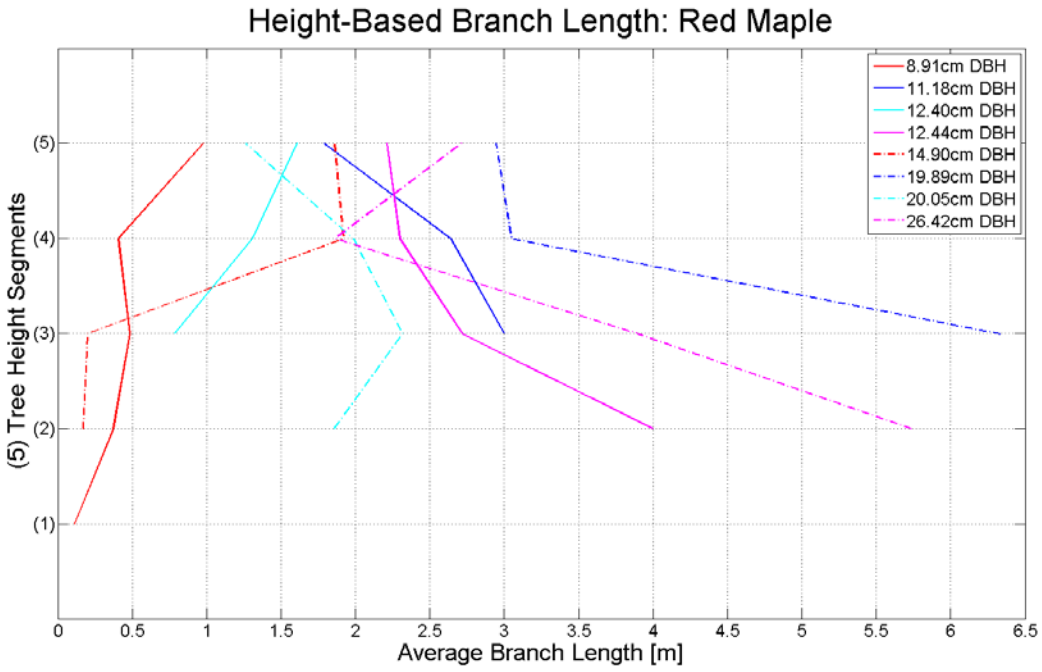

Figure 61: Height-Based Branch Length for Red Maple by Segment

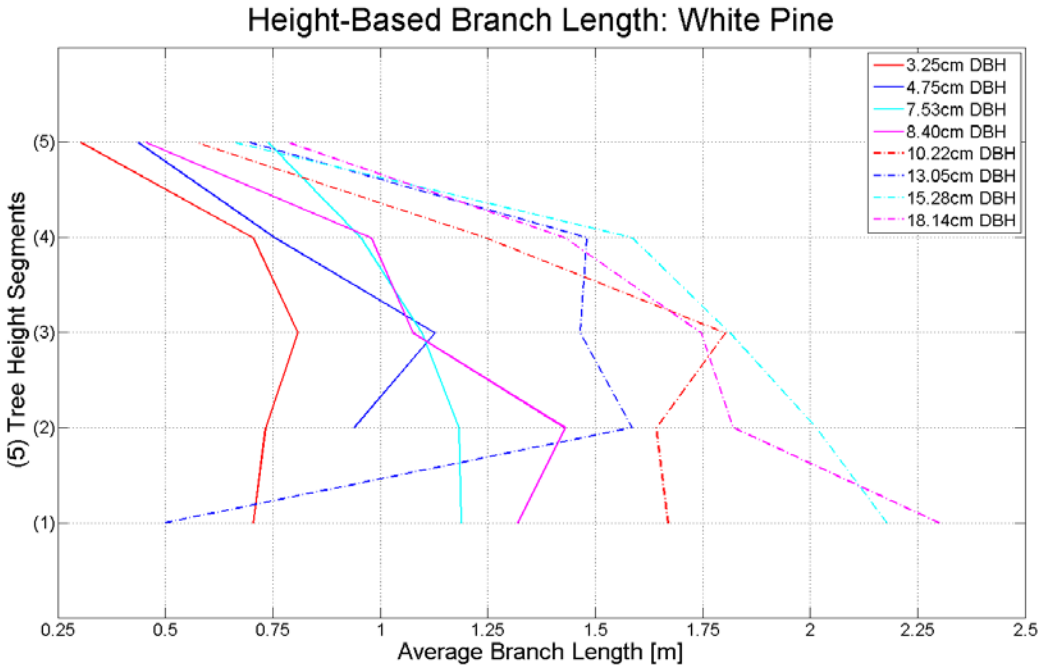

Figure 62: Height-Based Branch Length for White Pine by Segment

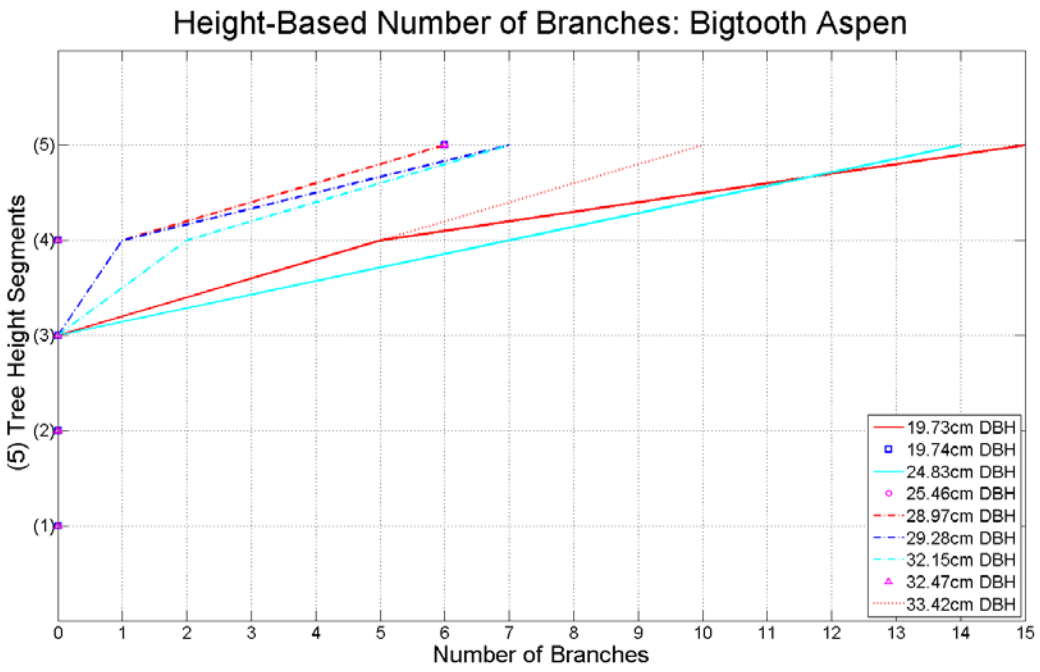

Figure 63: Height-Based Number of Branches for Bigtooth Aspen by Segment

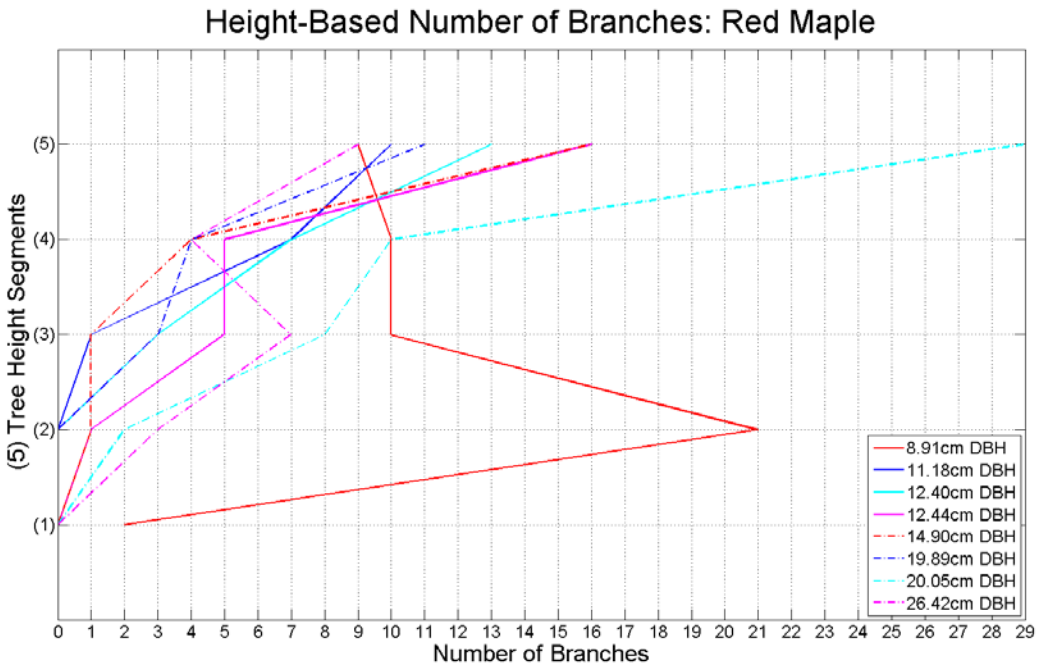

Figure 64: Height-Based Number of Branches for Red Maple by Segment

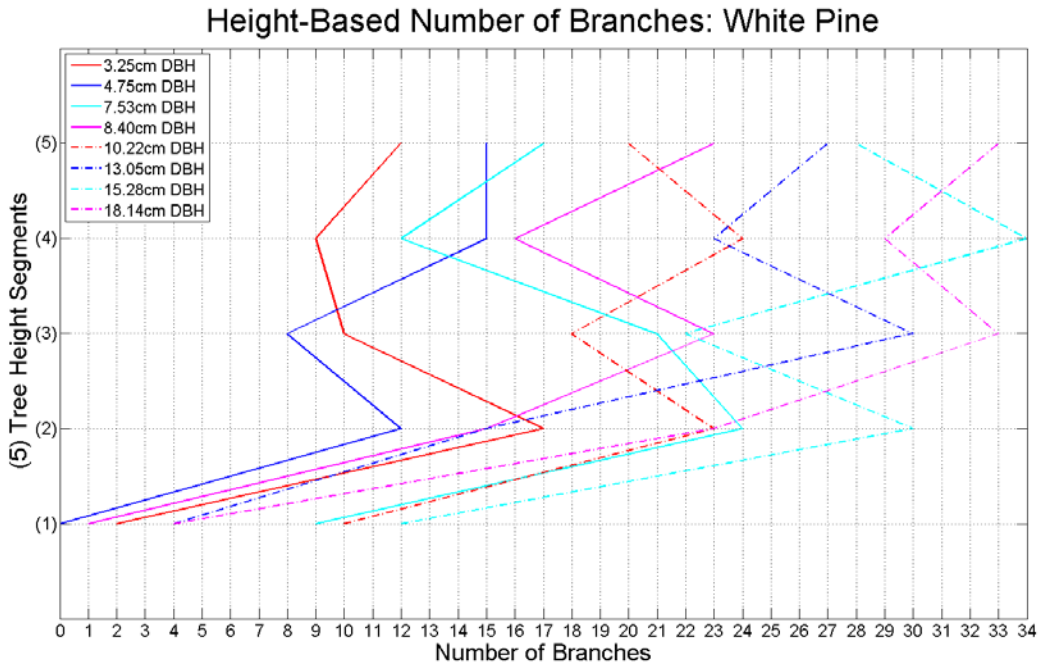

Figure 65: Height-Based Number of Branches for White Pine by Segment

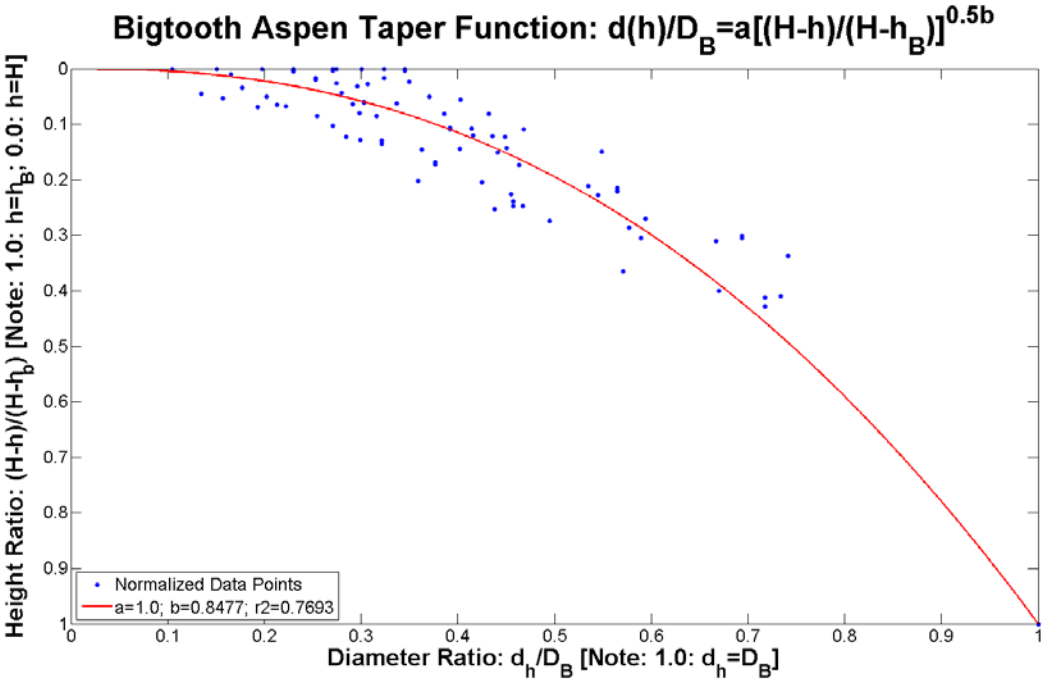

Figure 66: Taper Function Defining Bigtooth Aspen Branches and Stem

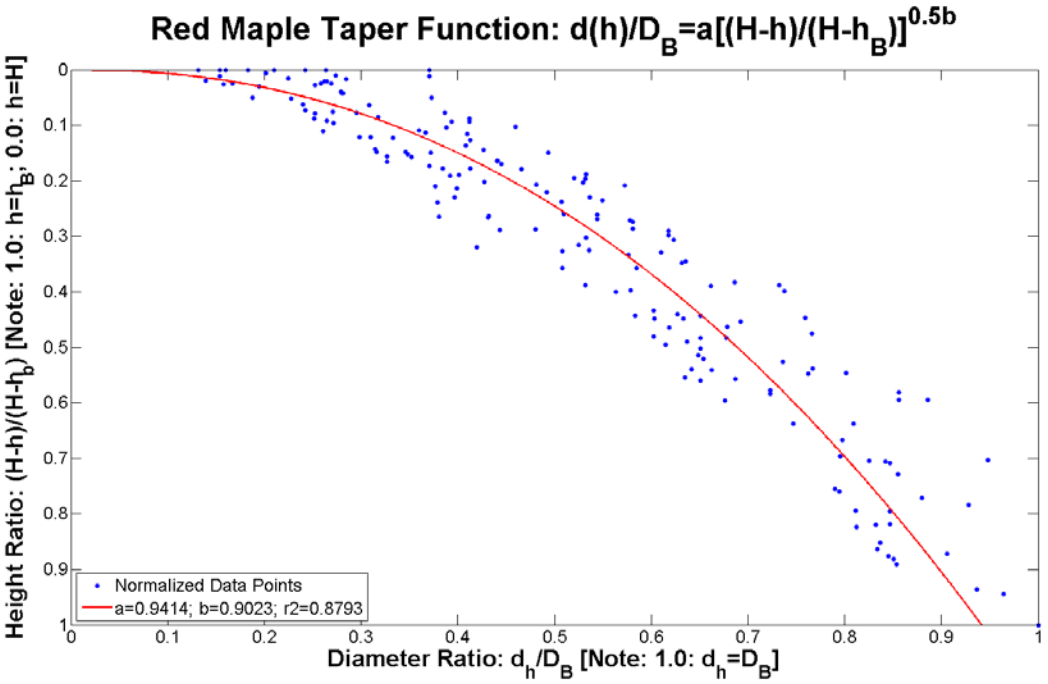

Figure 67: Taper Function Defining Red Maple Branches and Stem

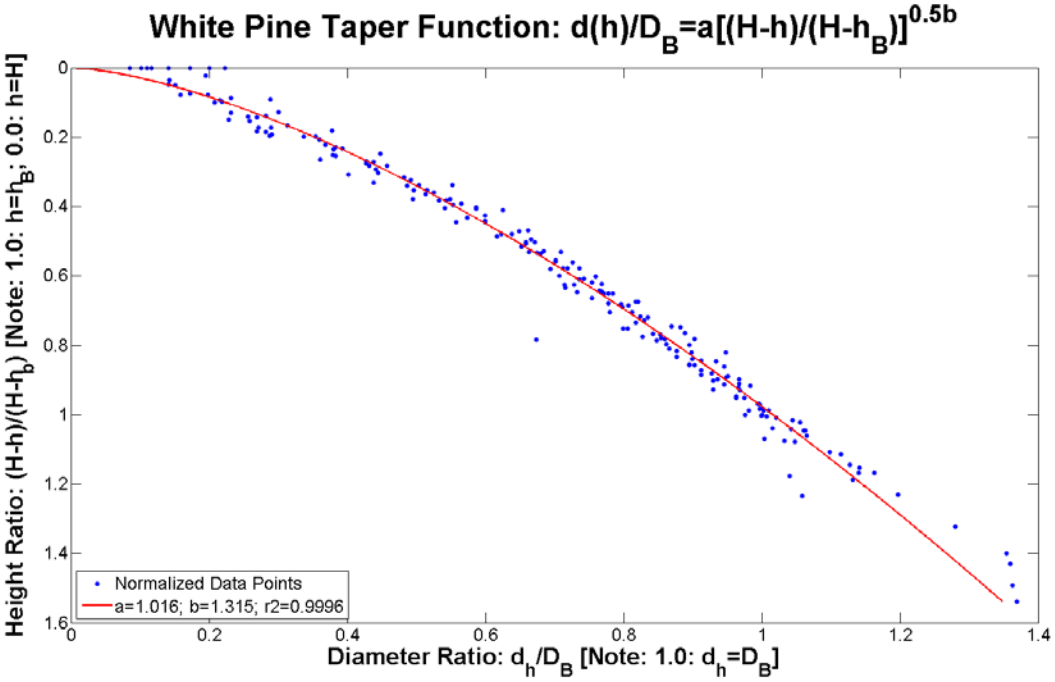

Figure 68: Taper Function Defining White Pine Branches and Stem

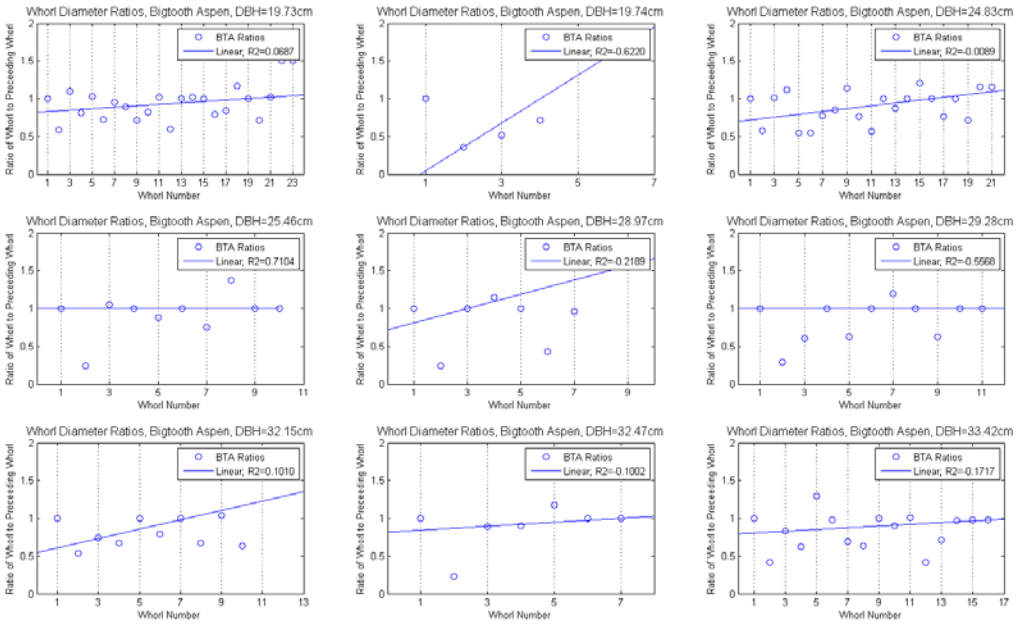

Figure 69: (9) Bigtooth Aspen Whorl Diameter Ratio Plots

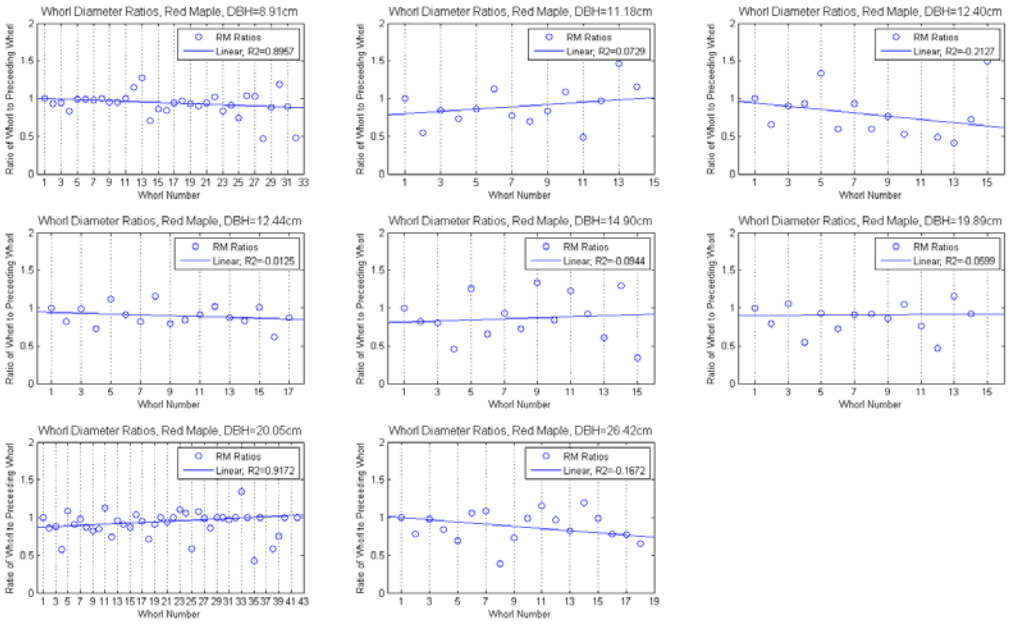

Figure 70: (8) Red Maple Whorl Diameter Ratio Plots

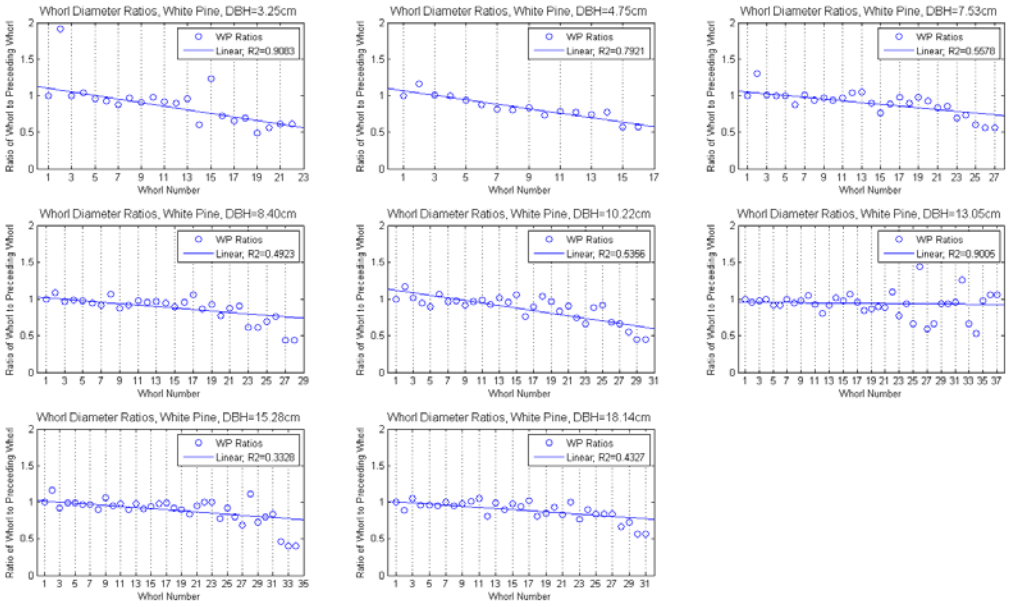

Figure 71: (8) White Pine Whorl Diameter Ratio Plots

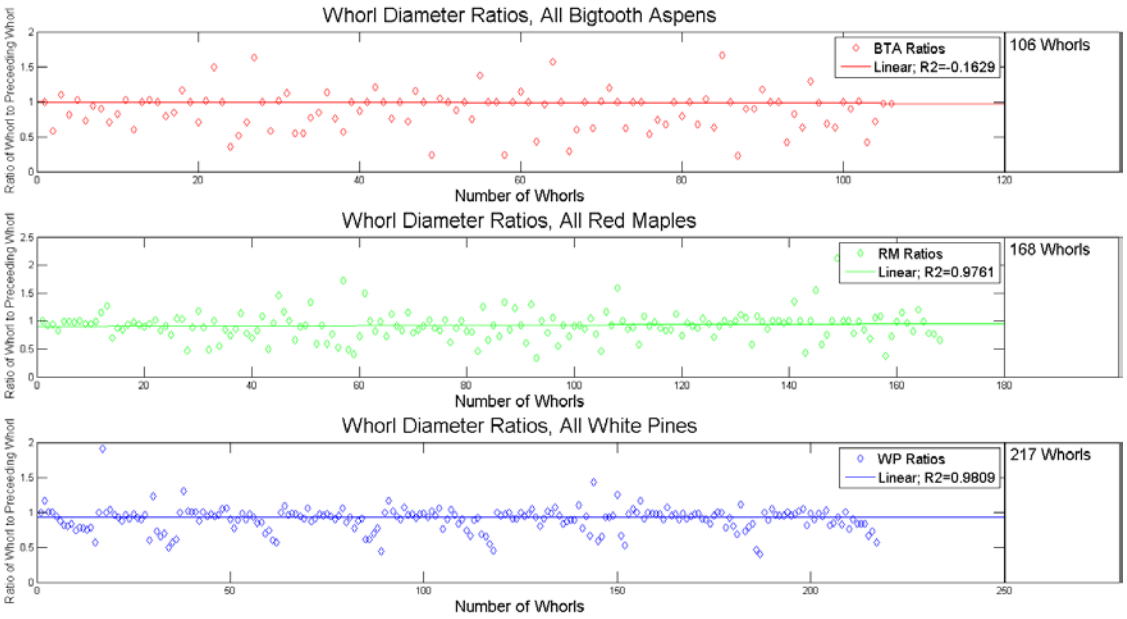

Figure 72: Summary Plots of All Whorl Diameter Ratios

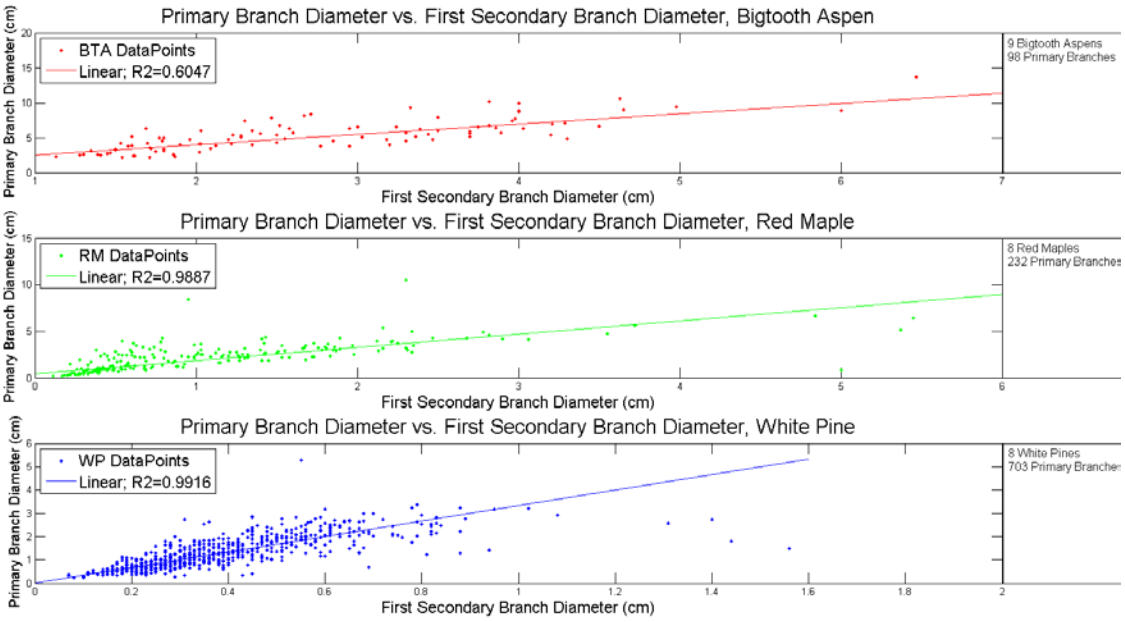

Figure 73: Summary Plots of Primary Branch Diameters versus First Secondary Branch Diameters

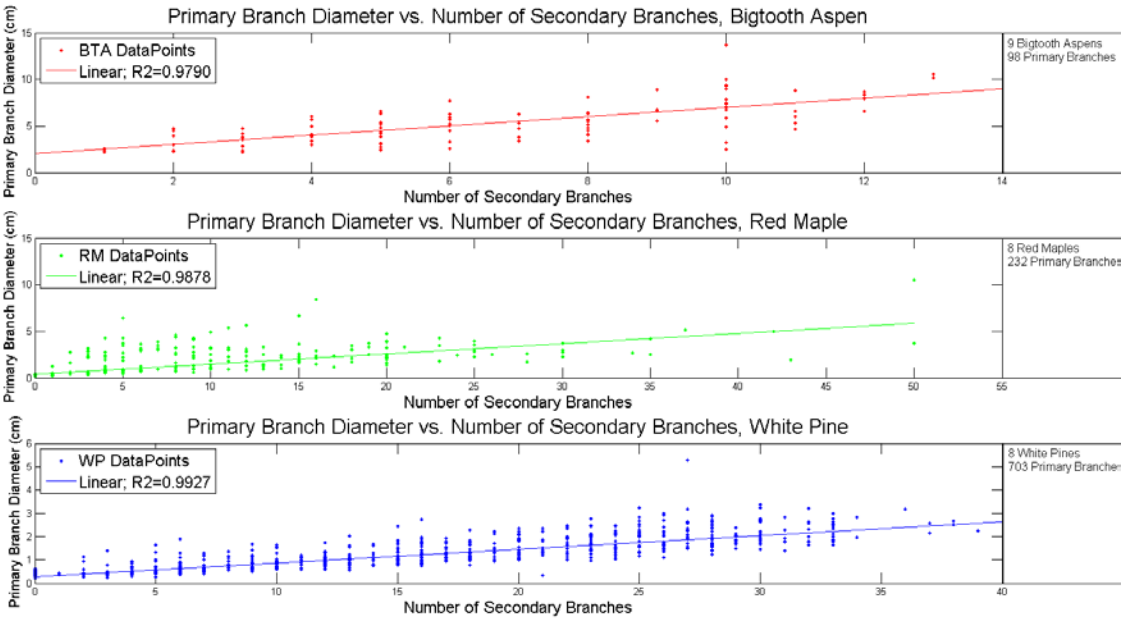

Figure 74: Summary Plots of Primary Branch Diameters versus Number of Secondary Branches

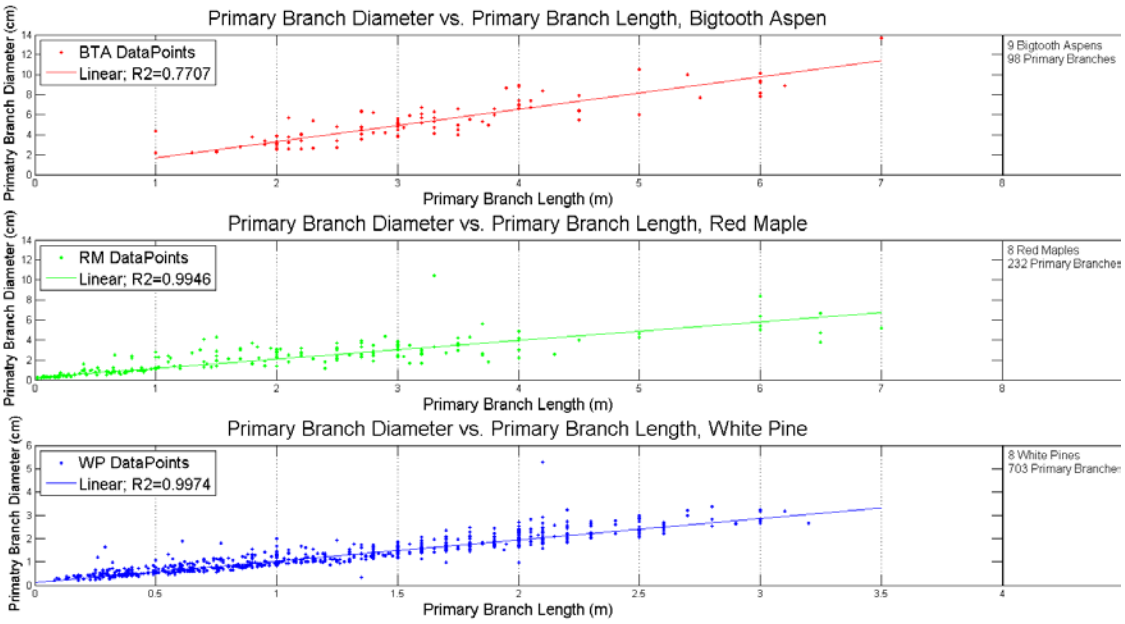

Figure 75: Summary Plots of Primary Branch Diameters versus Primary Branch Lengths

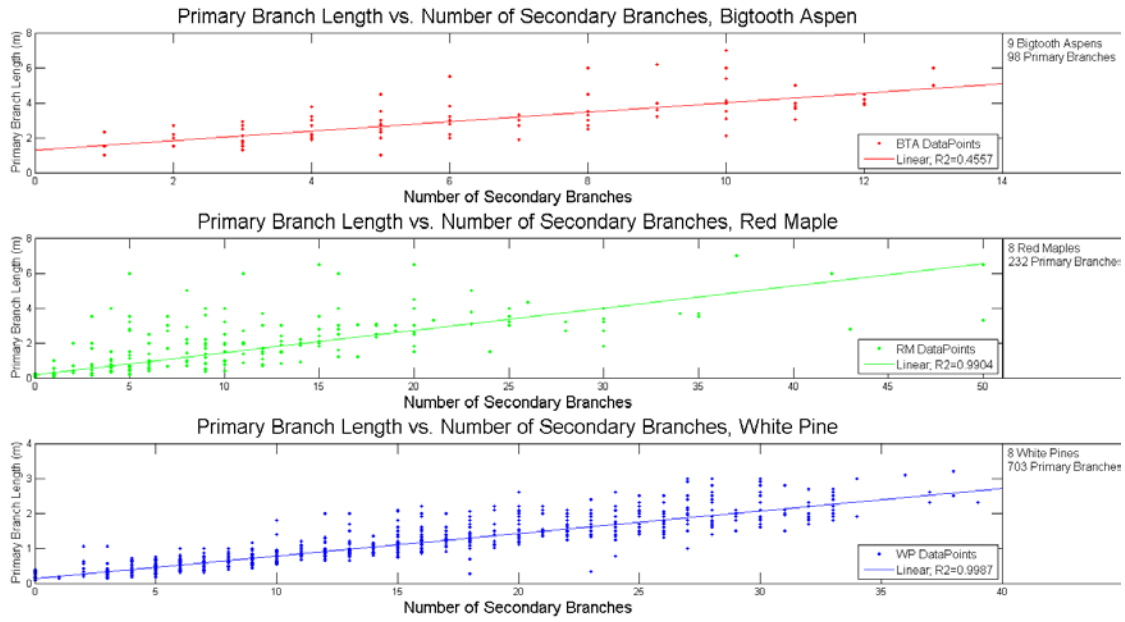

Figure 76: Summary Plots of Primary Branch Lengths versus Number of Secondary Branches

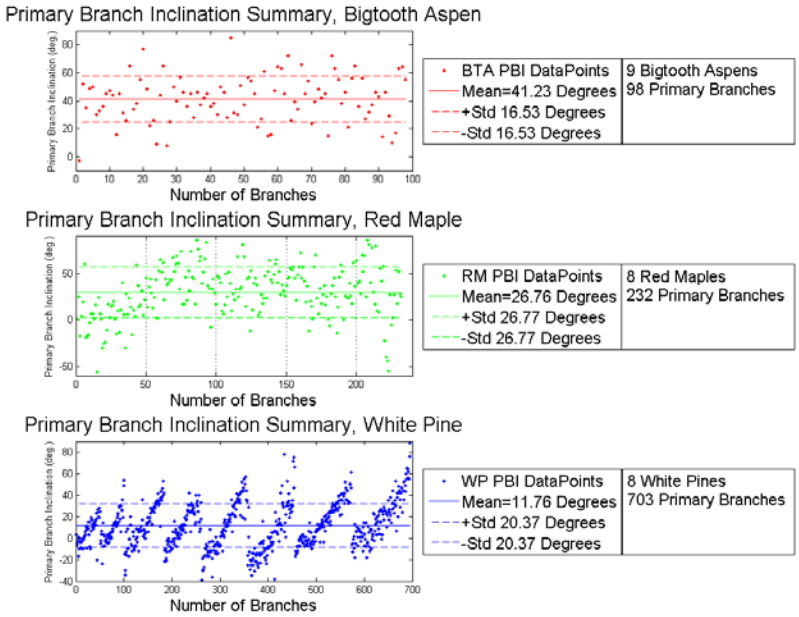

Figure 77: Summary Plots of Primary Branch Inclinations versus Number of Branches

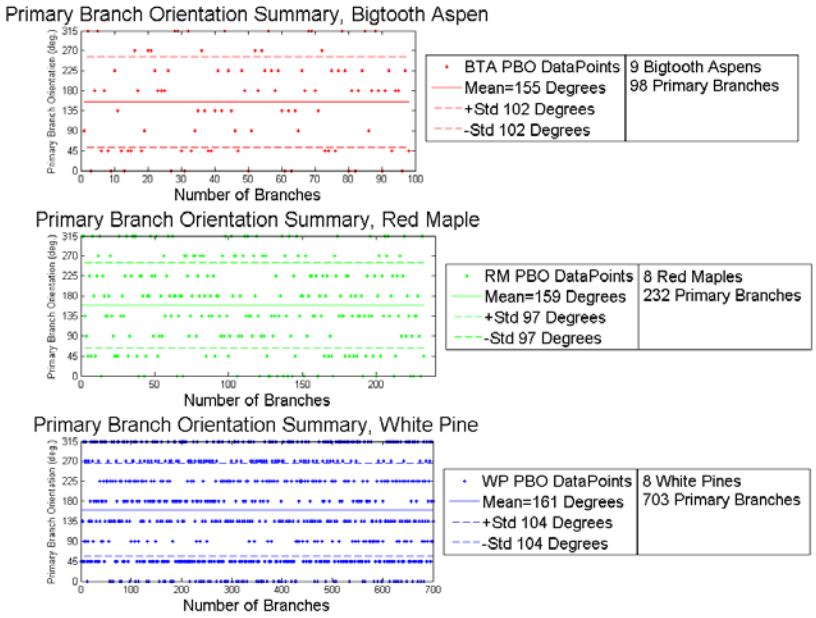

Figure 78: Summary Plots of Primary Branch Orientations versus Number of Branches

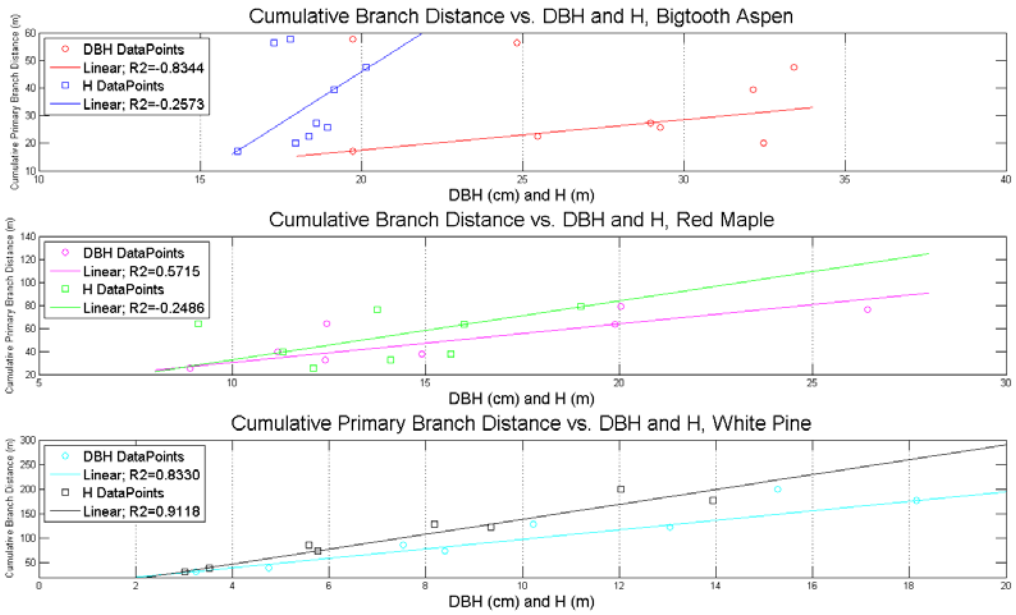

Figure 79: Summary Plots of Cumulative Branch Distances versus DBH's and H's

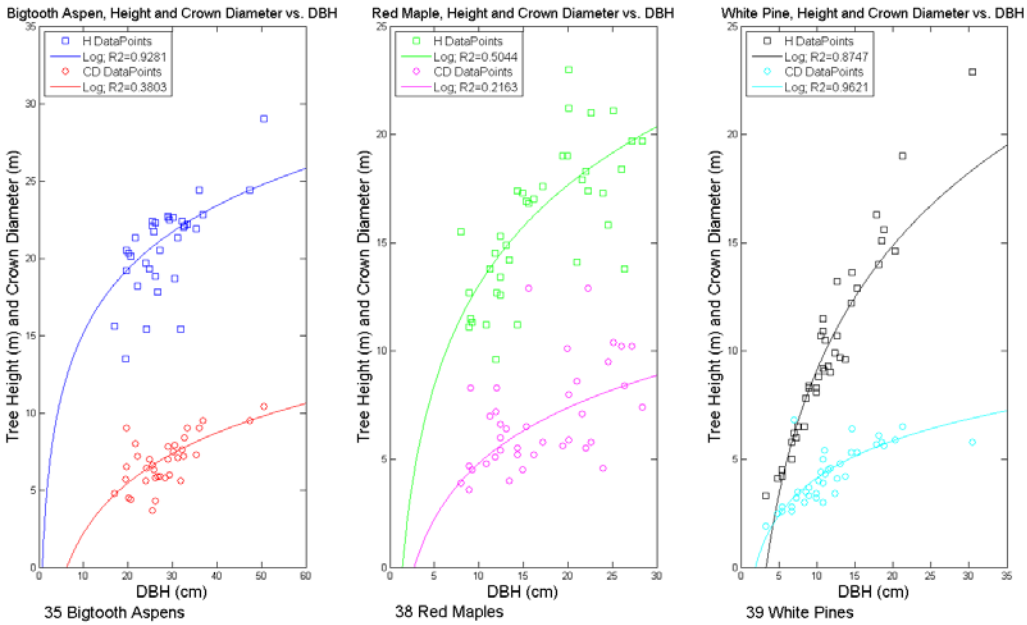

Figure 80: Species Specific Summary Plots of H's and CD's versus DBH's

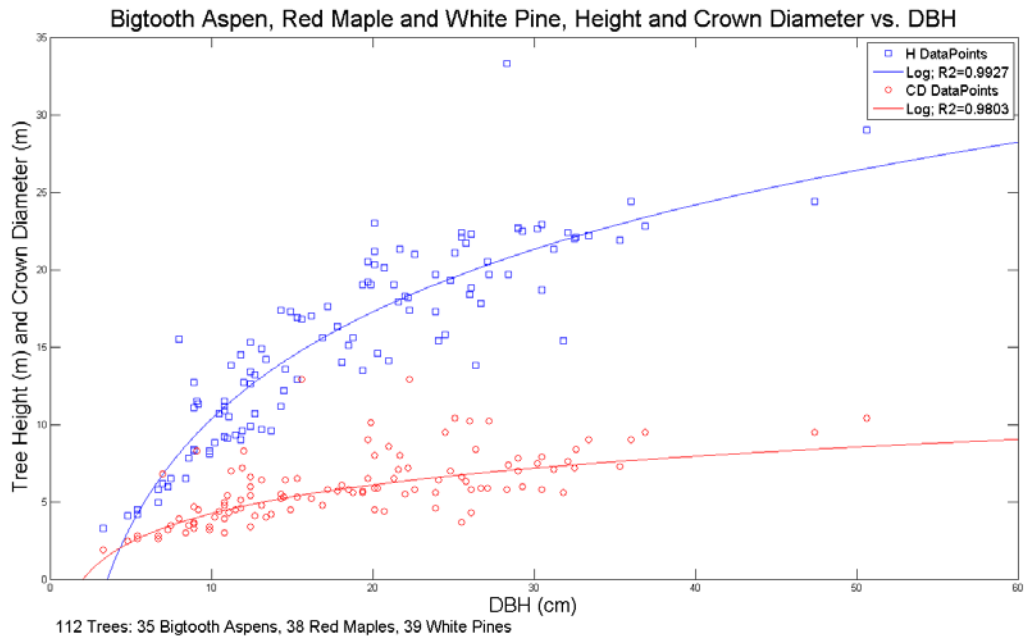

Figure 81: Combined Species Summary Plots of H's and CD's versus DBH's

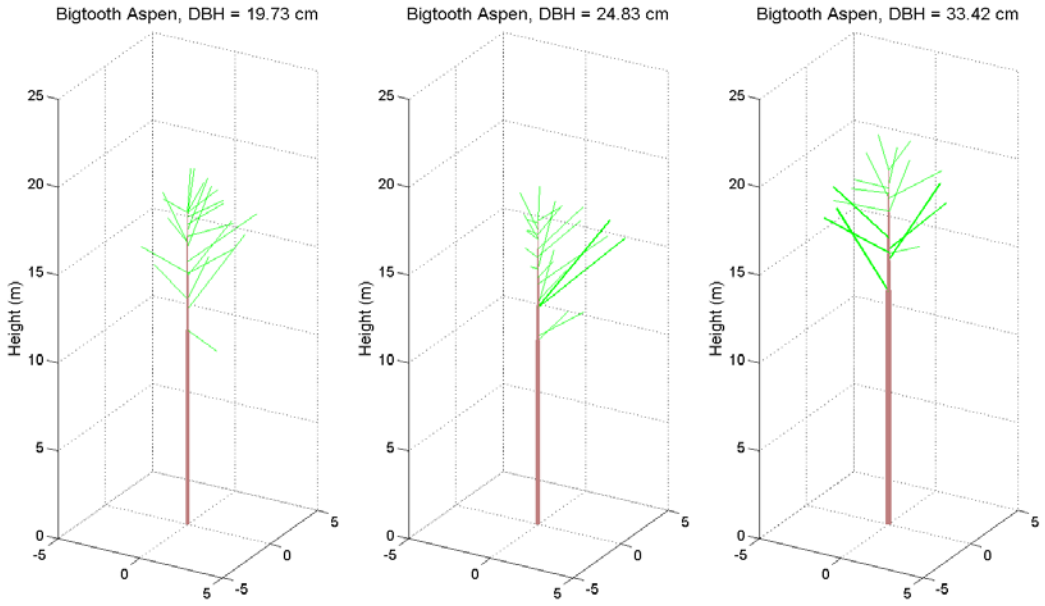

Figure 82: Bigtooth Aspen Pipe Model Plots

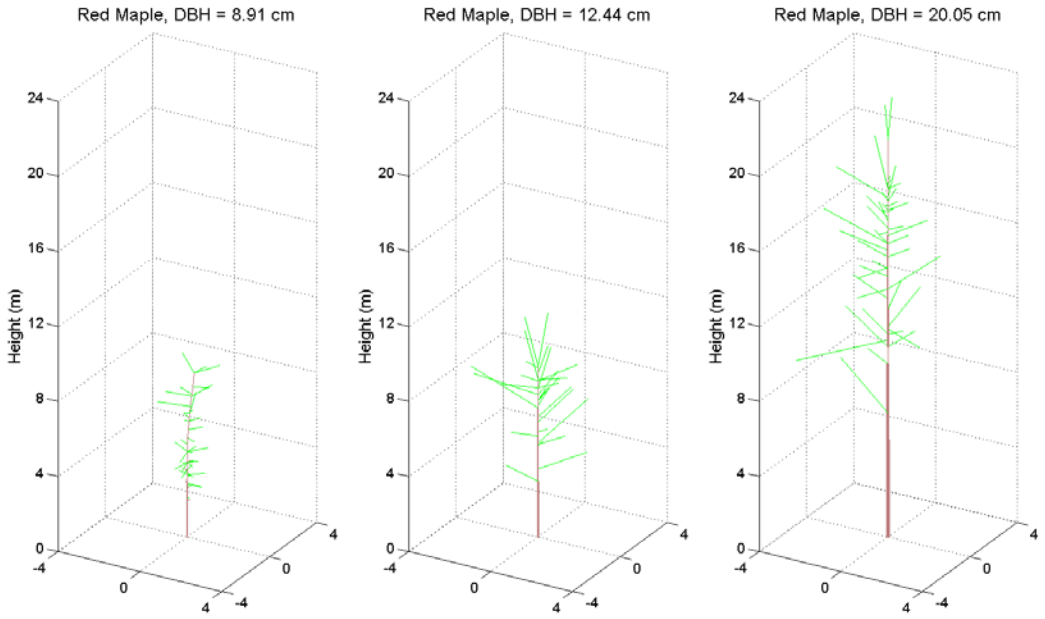

Figure 83: Red Maple Pipe Model Plots

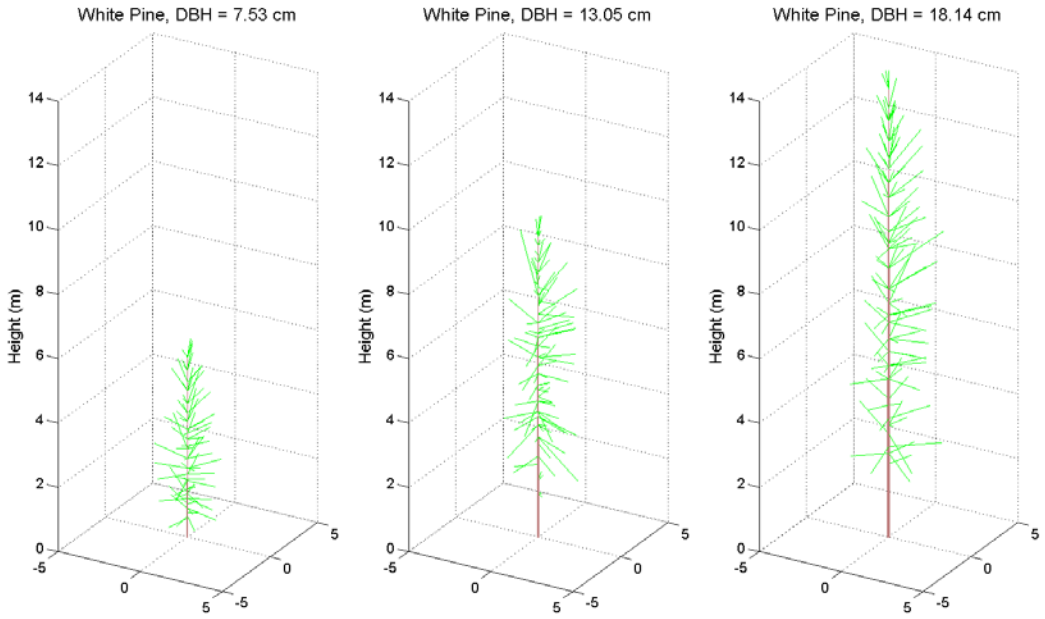

Figure 84: White Pine Pipe Model Plots

Table 2: Summary of Individual Trees Measured

Tree Summary									
Species	Bigtooth Aspen			Red Maple			White Pine		
Measurement	DBH (cm)	H (m)	CD (m)	DBH (cm)	H (m)	CD (m)	DBH (cm)	H (m)	CD (m)
Tree 1	19.73	20.50	9.0	8.91	12.65	4.7	3.25	3.29	1.9
Tree 2	19.74	16.19	6.5	11.18	13.80	7.0	4.75	4.13	2.5
Tree 3	24.83	19.30	7.0	12.40	15.29	6.6	7.53	6.48	3.5
Tree 4	25.46	22.38	6.6	12.44	12.62	6.0	8.40	6.53	3.0
Tree 5	28.97	18.64	7.0	14.90	17.30	4.5	10.22	8.79	4.0
Tree 6	29.28	22.45	6.0	19.89	19.00	10.1	13.05	9.65	4.8
Tree 7	32.15	22.40	7.6	20.05	21.20	8.0	15.28	12.87	5.3
Tree 8	32.47	17.99	7.2	26.42	13.75	8.4	18.14	13.97	6.1
Tree 9	33.42	22.15	9.0	NA	NA	NA	NA	NA	NA

Table 3: Taper Coefficients for All Trees

Taper Coefficients						
Species	Bigtooth Aspen		Red Maple		White Pine	
Variable	a	b	a	b	a	b
Tree 1	1.0000	0.7565	0.8899	0.7848	1.0360	1.3740
Tree 2	1.0000	0.6538	1.0000	1.0610	0.9623	1.2650
Tree 3	1.1000	1.1870	1.1350	1.0560	1.0250	1.1610
Tree 4	1.0000	0.6684	0.9129	0.6626	0.9904	1.2710
Tree 5	1.0000	0.8175	0.9611	0.8217	1.0400	1.4170
Tree 6	1.0000	0.9566	1.0000	0.8054	1.0010	1.3420
Tree 7	1.0000	0.7825	0.9662	1.1500	1.0150	1.2940
Tree 8	1.0000	0.8669	1.0910	1.2070	1.0000	1.2770
Tree 9	1.0000	0.9753	NA	NA	NA	NA
Mean	1.01	0.85	0.99	0.94	1.01	1.30
Std. Dev.	0.03	0.17	0.08	0.20	0.03	0.08